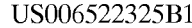
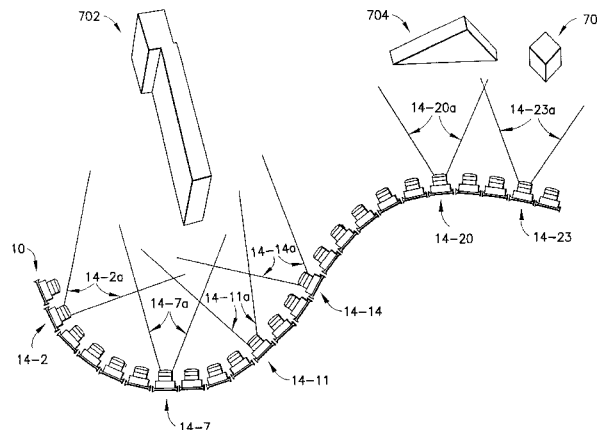


Exhibit B



(10) **Patent No.:** US 6,522,325 B1
(45) **Date of Patent:** Feb. 18, 2003

- ### 30 Claims, 13 Drawing Sheets



U.S. PATENT DOCUMENTS

5,488,674 A	1/1996	Burt et al.	382/284
5,495,576 A	2/1996	Ritchey	345/420
5,497,188 A	3/1996	Kaye	348/36
5,562,572 A	10/1996	Carmein	482/4
5,566,251 A	10/1996	Hanna et al.	382/284
5,581,629 A	12/1996	Hanna et al.	382/103
5,585,858 A	12/1996	Harper et al.	382/103
5,598,208 A	1/1997	McClintock	348/159
5,600,368 A	2/1997	Matthews, III	348/143
5,604,529 A	2/1997	Kuga et al.	348/46
5,616,912 A	4/1997	Robinson et al.	250/201.1
5,629,988 A	5/1997	Burt et al.	382/276
5,632,007 A	5/1997	Freeman	706/59
5,644,694 A	7/1997	Appleton	345/474
5,649,032 A	7/1997	Burt et al.	382/284
5,659,323 A	8/1997	Taylor	348/159
5,682,196 A	10/1997	Freeman	725/139
5,686,957 A	11/1997	Baker	348/36
5,703,961 A	12/1997	Rogina et al.	382/154
5,706,416 A	1/1998	Mann et al.	345/427
5,708,469 A	1/1998	Herzberg	348/39
5,724,091 A	3/1998	Freeman et al.	725/138
5,729,471 A	3/1998	Jain et al.	725/131
5,745,305 A	4/1998	Nalwa	359/725
5,850,352 A	12/1998	Moezzi et al.	345/419
5,900,849 A	5/1999	Gallery	345/8
5,963,664 A	10/1999	Kumer et al.	382/154
5,999,641 A	12/1999	Miller et al.	382/154
6,020,931 A	2/2000	Bilbrey et al.	348/584
6,084,979 A	7/2000	Kanade et al.	382/154
6,151,009 A	11/2000	Kanade et al.	345/641
6,154,251 A	11/2000	Taylor	348/159
6,208,379 B1	3/2001	Oya et al.	348/213

OTHER PUBLICATIONS

Carnegie Mellon University Virtualized Reality, <http://www.cs.cmu.edu/af/cs/project/VirtualizedR/www/explain.html>.
 Carnegie Mellon University, <http://www.cs.cmu.edu/afs/cs/project/VirtualizedR/www/main.html>.
 Carnegie Mellon University, http://www.cs.cmu.edu/afs/cs/project/VirtualR/www/3manball_new.html.
 FullView Technologies, <http://www.fullview.com/about.html>.
 FullView Technologies, <http://www.fullview.com/technology.html>.
 FullView Technologies, <http://fullview.com/gallery.html>.
 Hipbone, <http://www.hipbone.com/solutions/>.

iMove/Infinite Pictures Inc./SmoothMove, <http://www.smoothmove.com>.
 iMove, <http://www.smoothmove.com/imovebeta/partners/index.html>.
 iMove, <http://smoothmove.com/imovebeta/partners/index.html>.
 IPIX/Interactive Pictures Corp., <http://www.ipix.com>.
 IPIX, <http://ipix.com/products/fmproducts.html>.
 Kumar, R., et al., "Detecting Lesions In Magnetic Resonance Breast Scans," SPIE, Oct. 1995.
 Multicam Clips, <http://www.reelefx.com/Multicam/Mmain.htm>.
 Multicam/Index, <http://www.reelefx.com/Multicam/Mtechnical.html>.
 PAWS—at-a-glance, <http://www.hitpaws.com/glance.html>.
 PAWS, <http://www.hitpaws.com/company.html>.
 PAWS, <http://www.hitpaws.com/faq.html>.
 Perceptual Robotics Inc., <http://www.perceptualrobotics.com/about/>.
 Perceptual Robotics Inc., <http://www.perceptualrobotics.com/products/>.
 Polycamera, <http://www.cs.columbia.edu/CAVE/research/demos>.
 Snell&Wilcox, <http://www.snell.co.uk/internet/aboutus/indexd.html>.
 Snell & Wilcom Matrix press release/Jun. 1999, <http://www.snell.co.uk/internet/pressindexd.html>.
 Stanford University, Light field rendering <http://graphics.stanford.edu/projects/lightfield/>.
 Timetrack, virtual camera movement, <http://www.virtualcamera.com/invention.html>.
 Timetrack, <http://www.virtualcamera.com/cameras.html>.
 Zaxel, 3D Replay, 2001.
 Zaxel, Virtual Viewpoint—Remote Collaboration, 2001.
 Zaxel, Virtual Viewpoint—Sports and Entertainment, 2001.
 Shenchang Eric Chen, "QuickTime VR—An Image-Based Approach to Virtual Environment Navigation," Apple Computer, Inc., pp. 1–10, Sep./1995.
 Kelly et al., "An Architecture for Multiple Perspective Interactive Video," pp., 1–20, 1995.
 Jasnoch et al., "Shared 3D Environment Within A Virtual Prototyping Environment," pp. 274–279, 1996.
 Leigh et al., "Multi-Perspective Collaborative Design in Persistent Networked Virtual Environments," pp. 1–13, 1996.

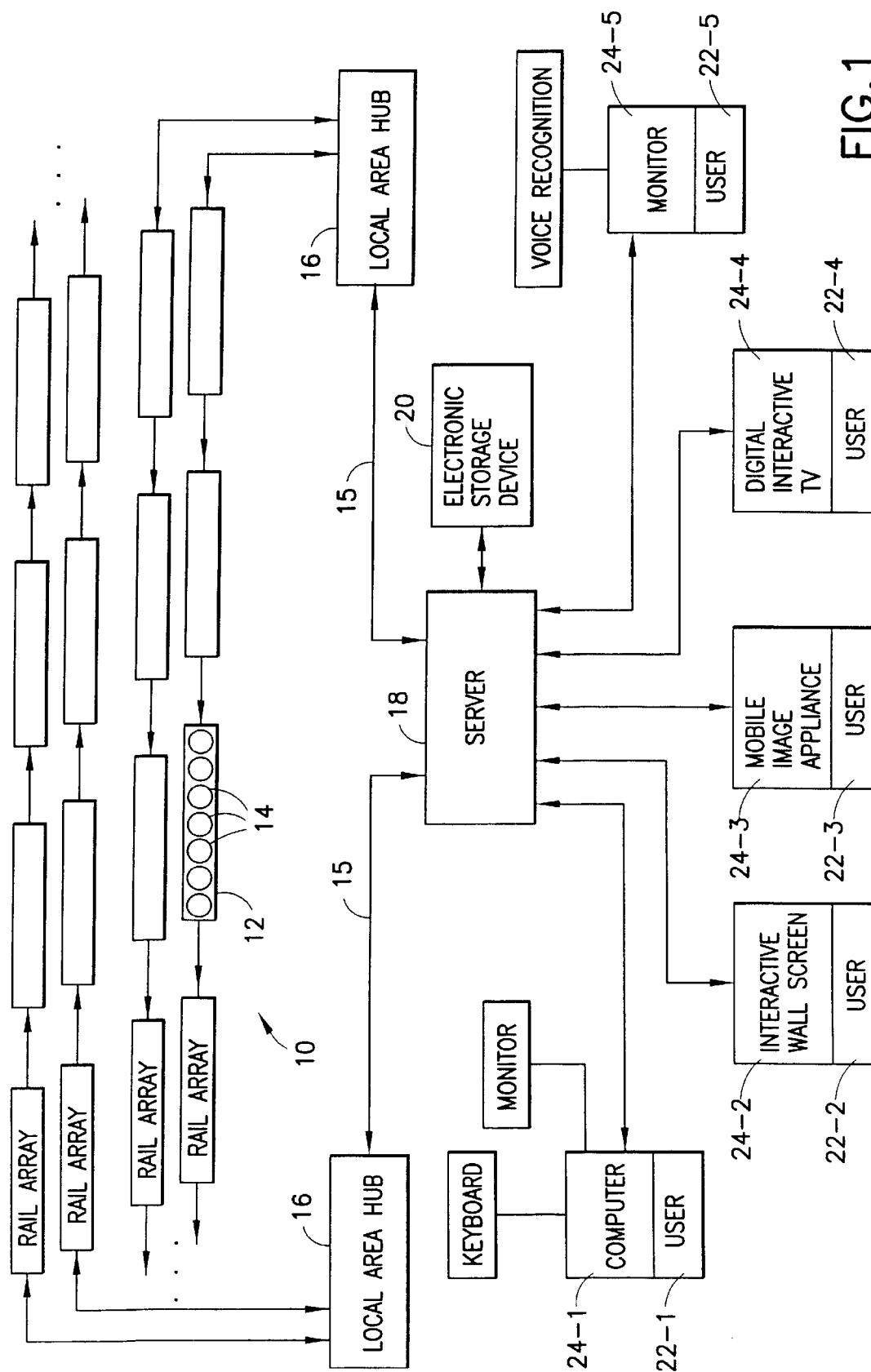


FIG. 1

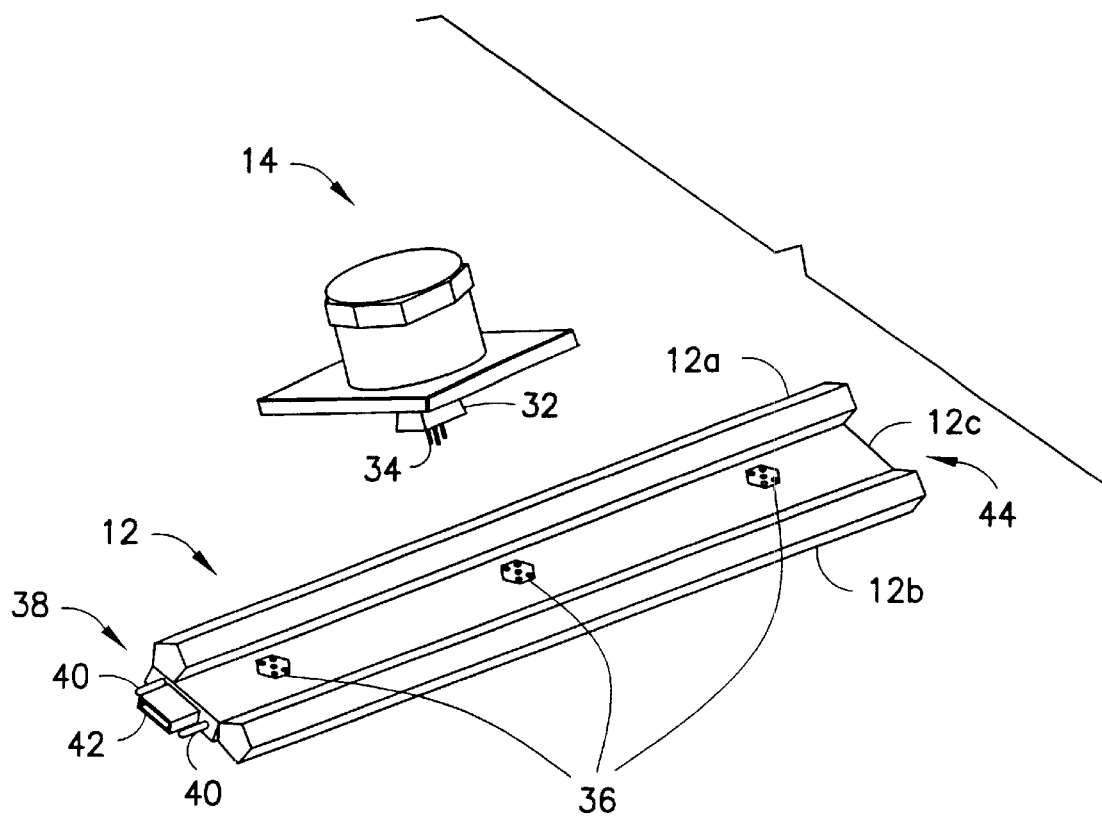


FIG.2a

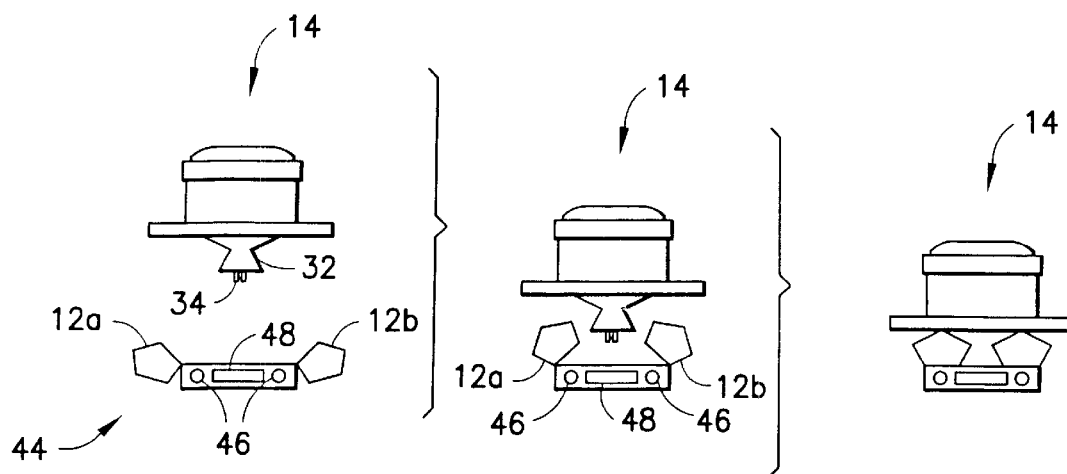


FIG. 2b

FIG. 2c

FIG. 2d

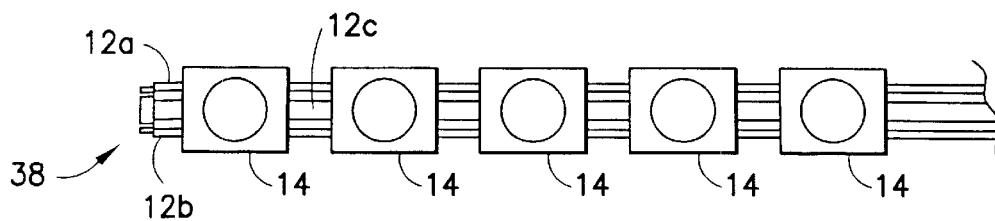
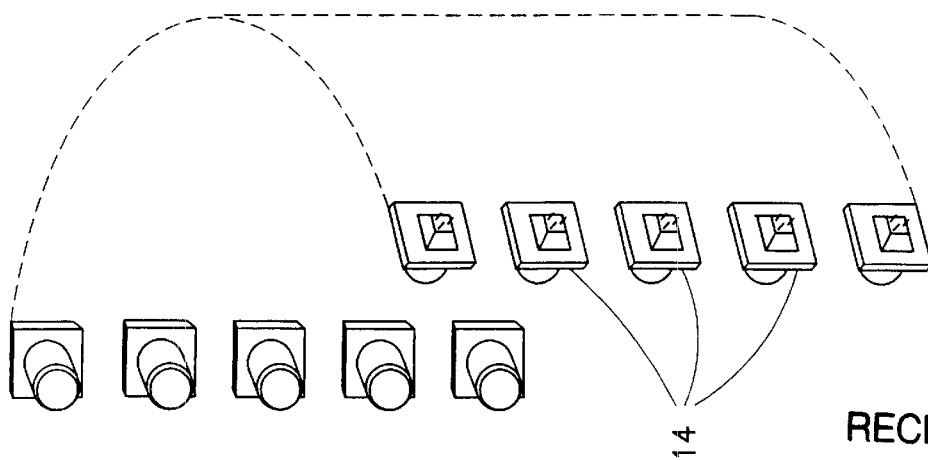


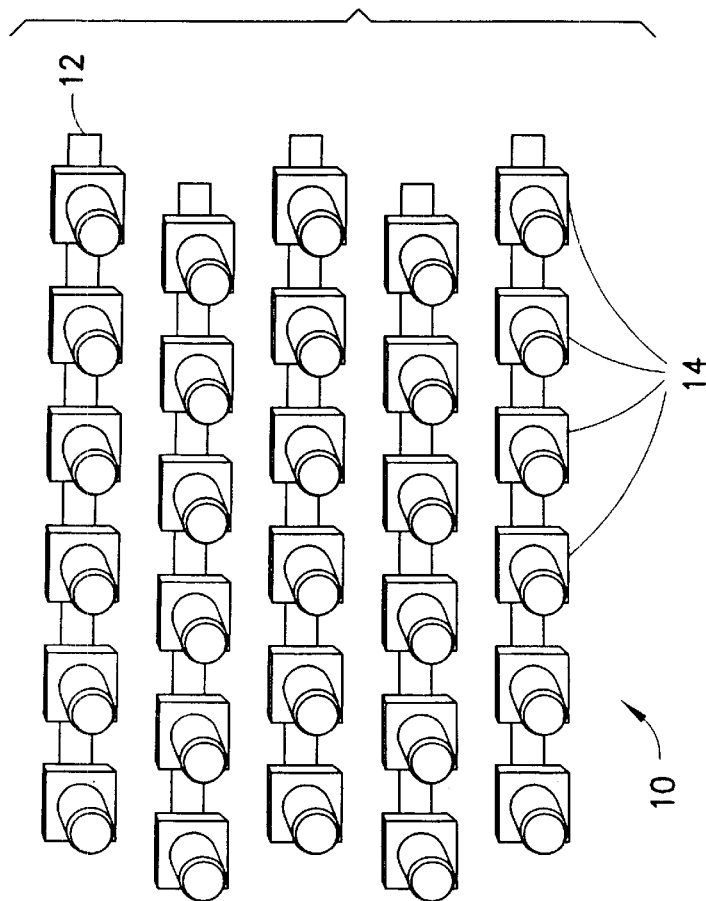
FIG. 2e



RECEIVED

JUL 26 2002

Technology Center 2100



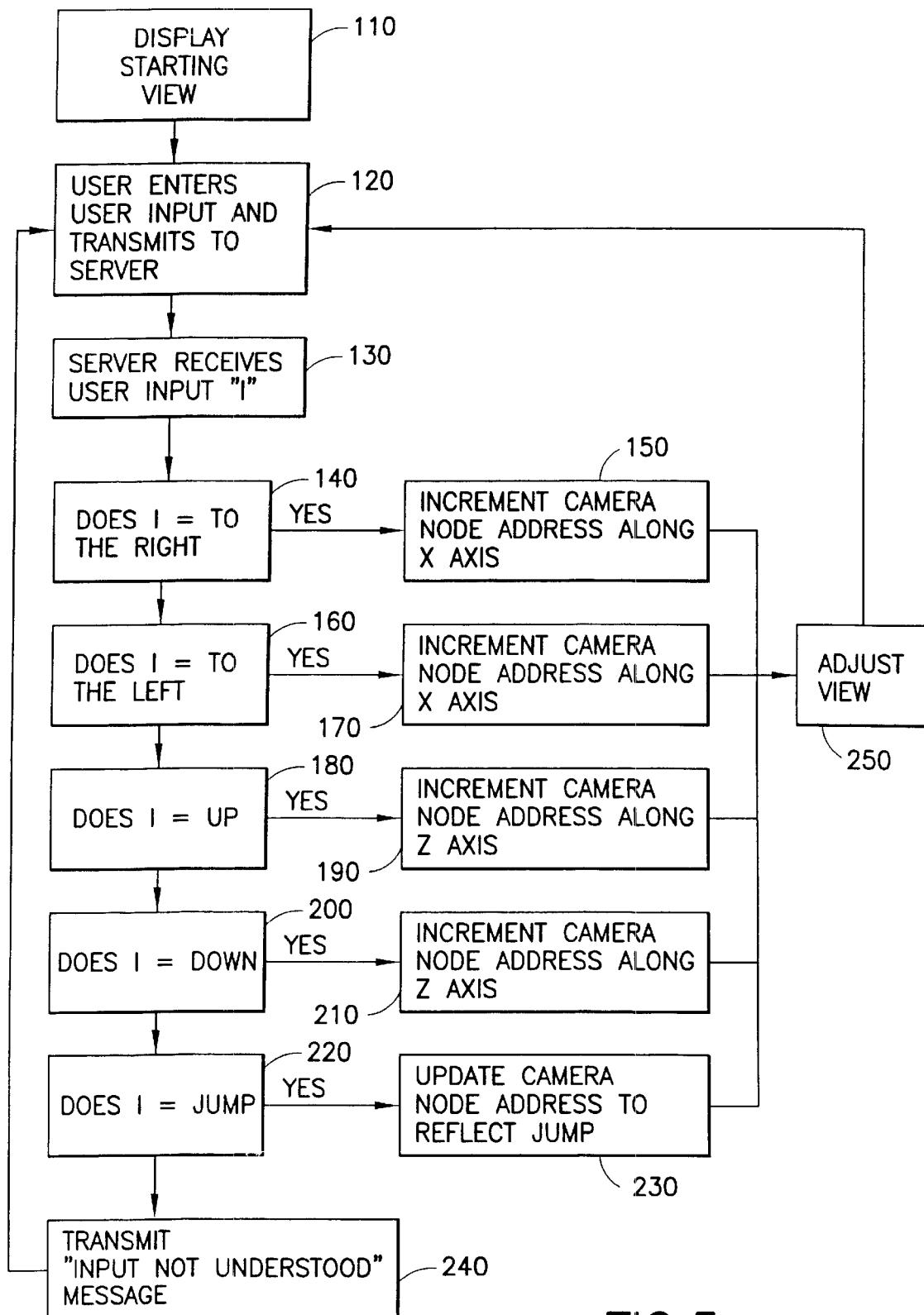


FIG.5

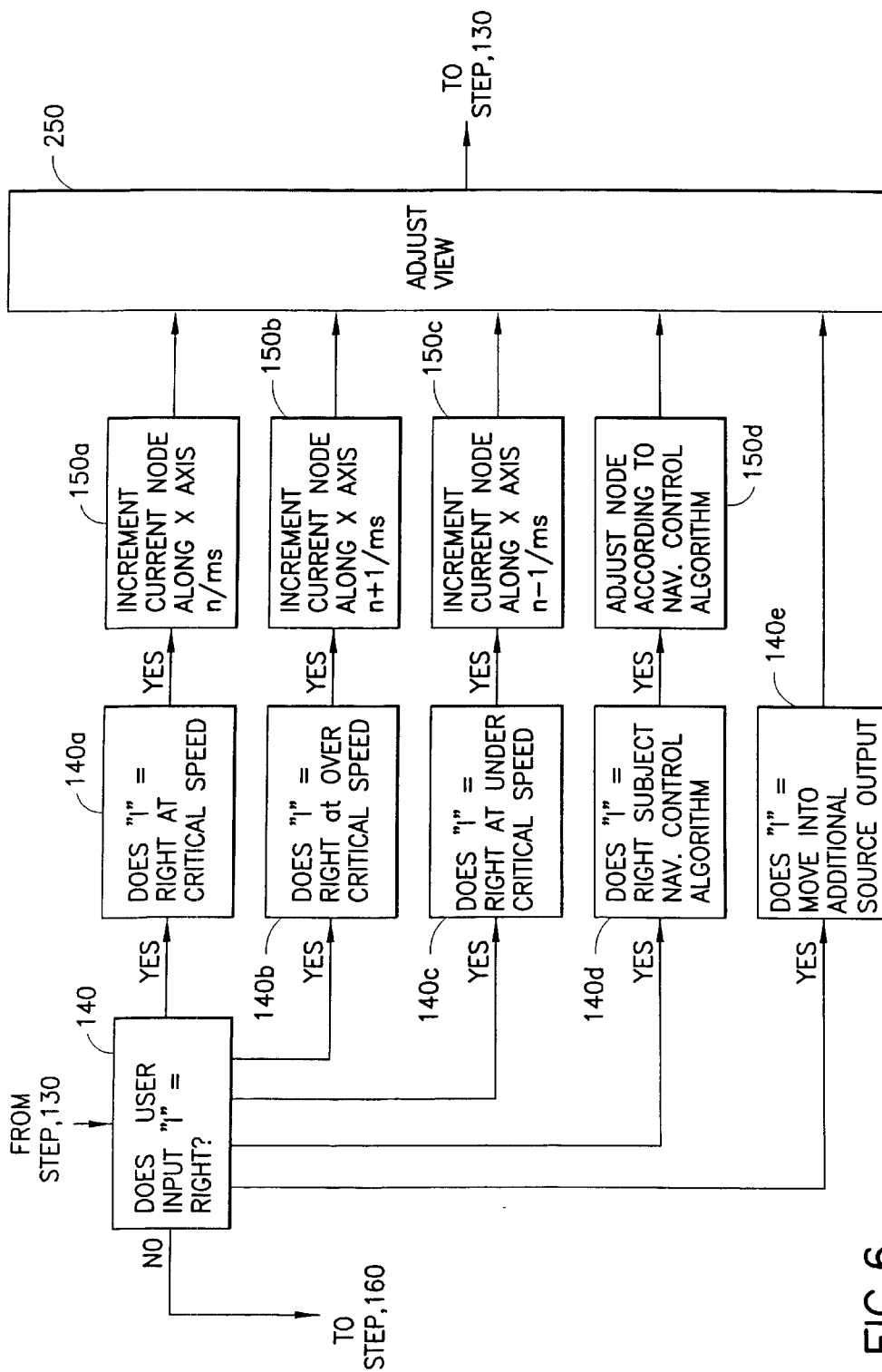


FIG. 6

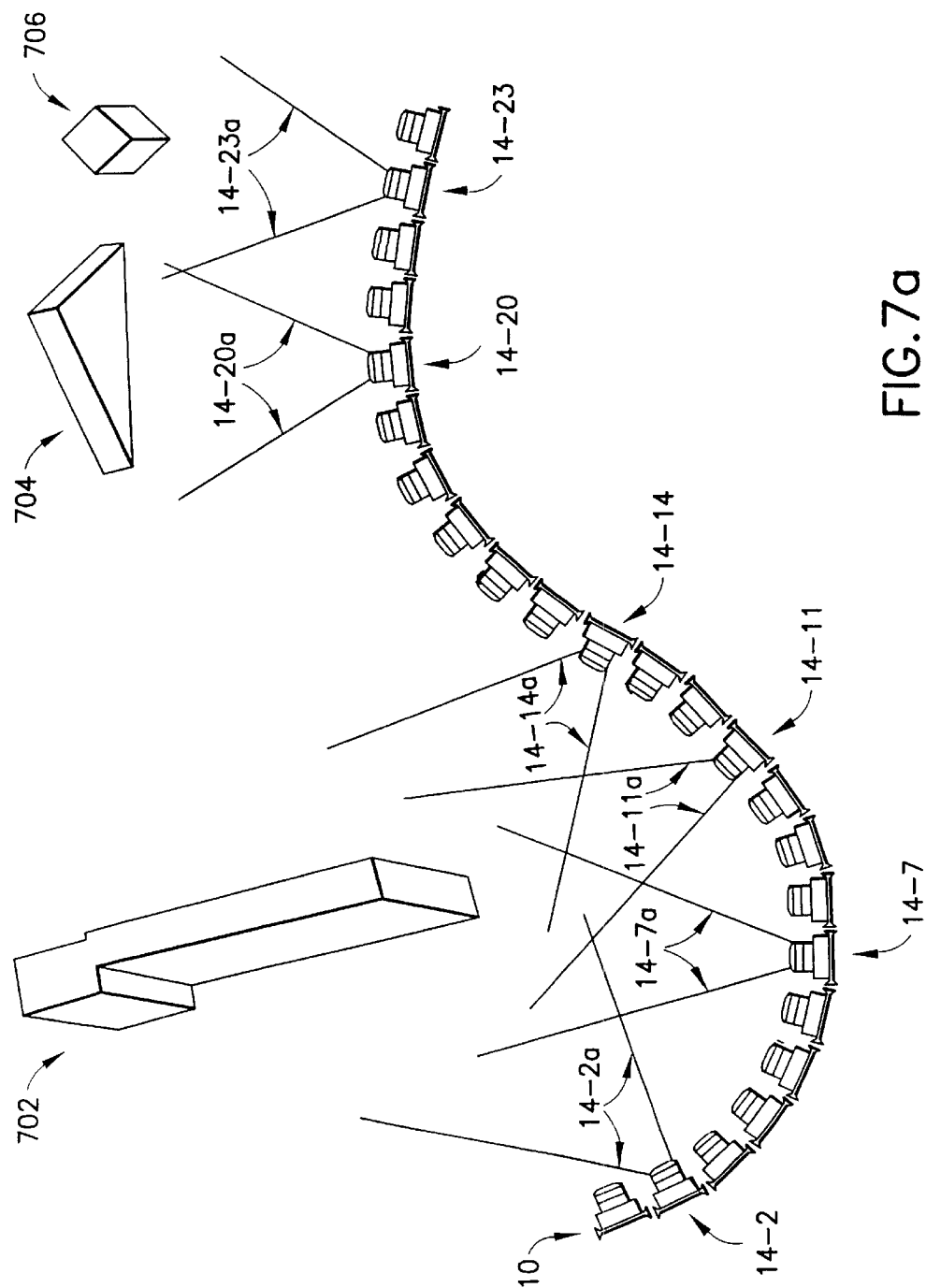


FIG. 7a

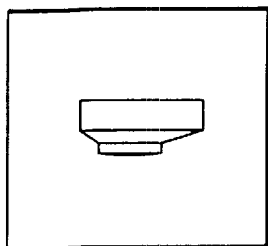


FIG. 7b

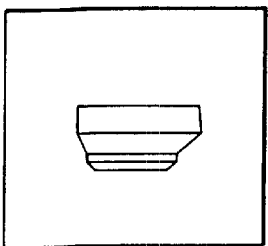


FIG. 7c

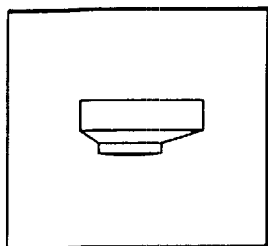


FIG. 7d

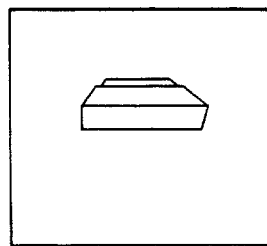


FIG. 7e

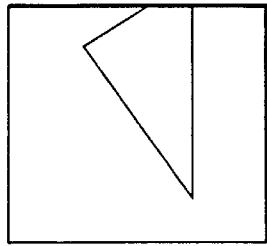


FIG. 7f

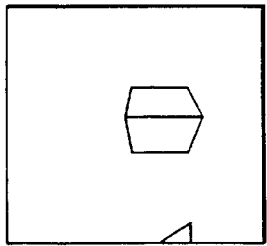


FIG. 7g

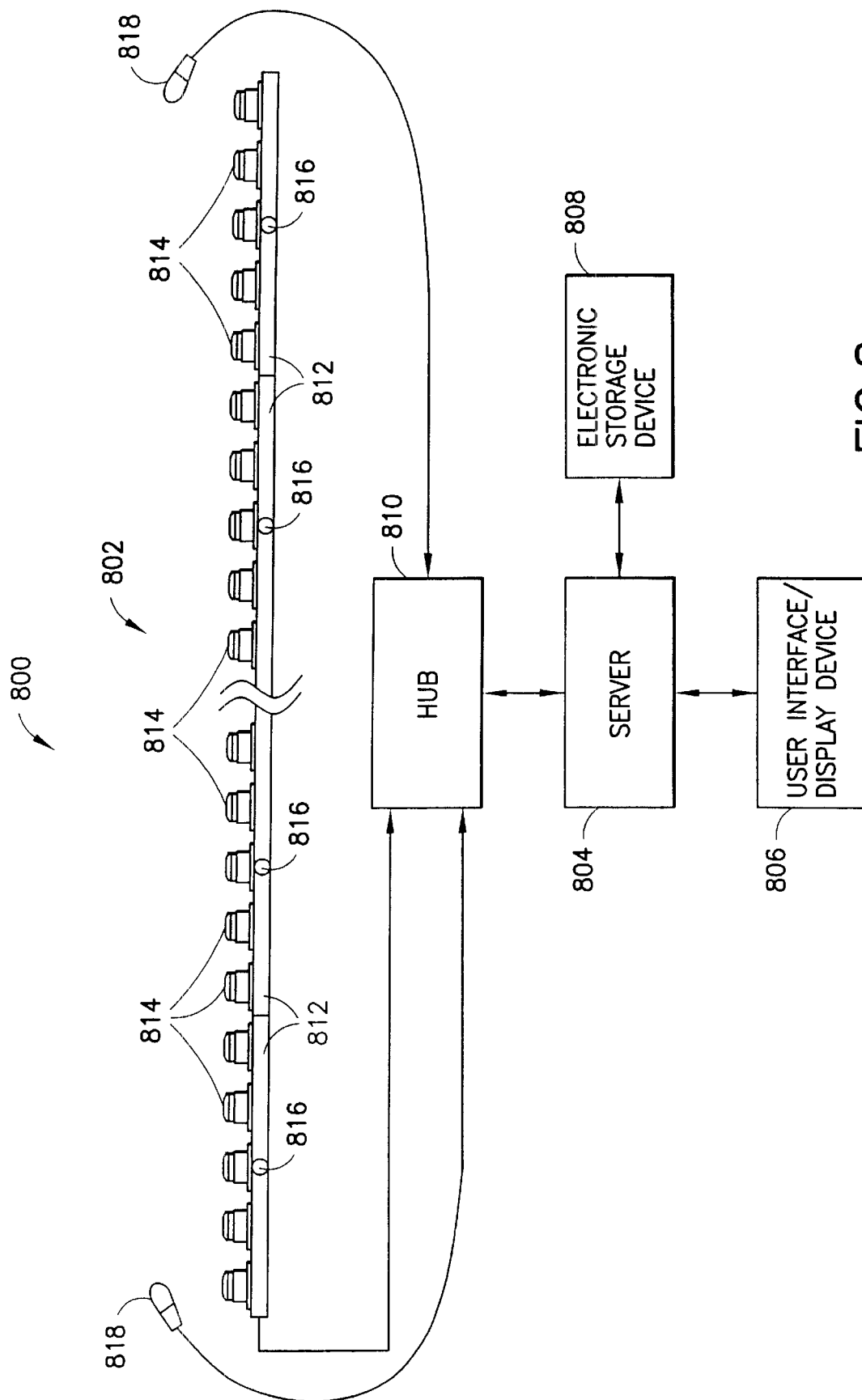
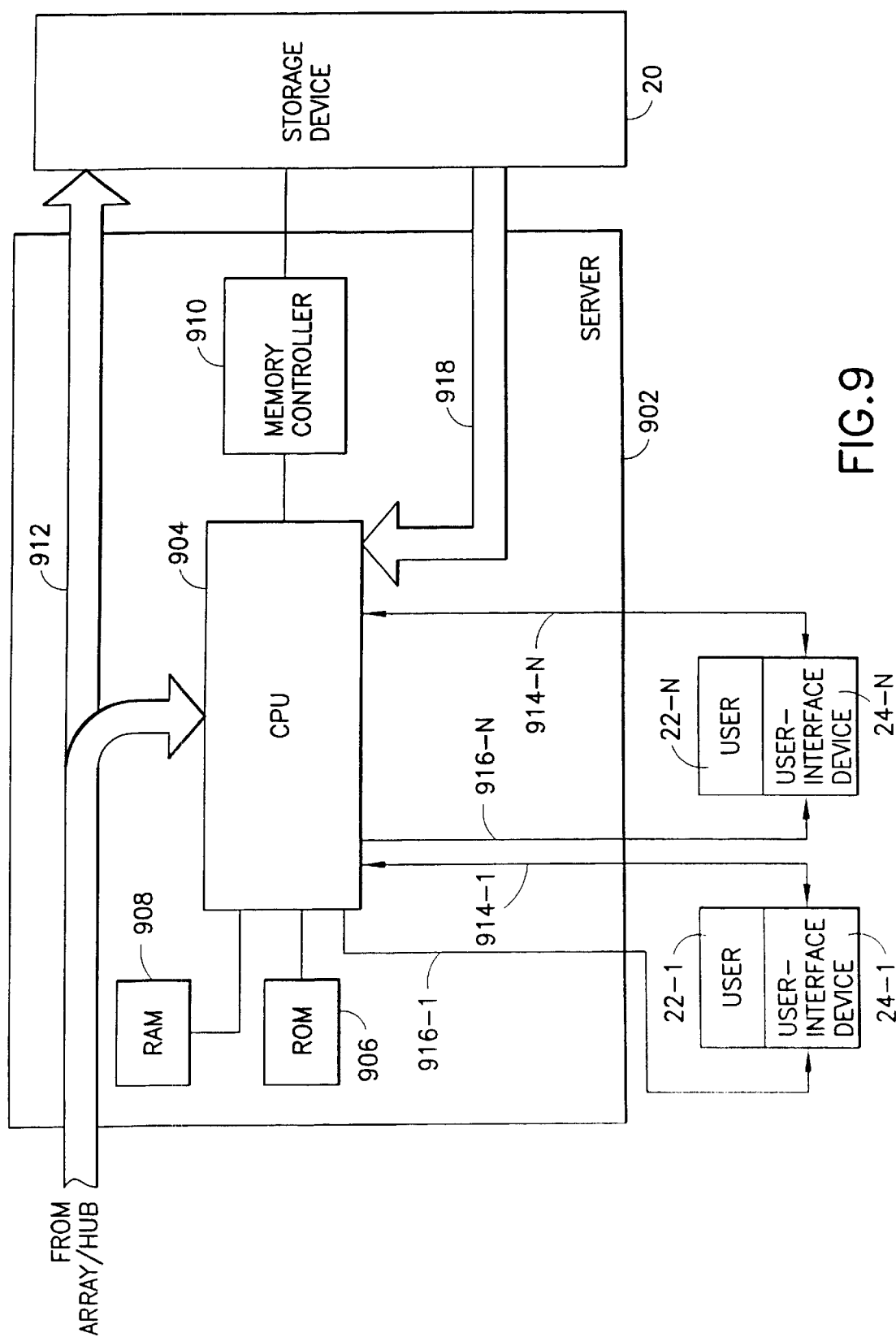


FIG. 8



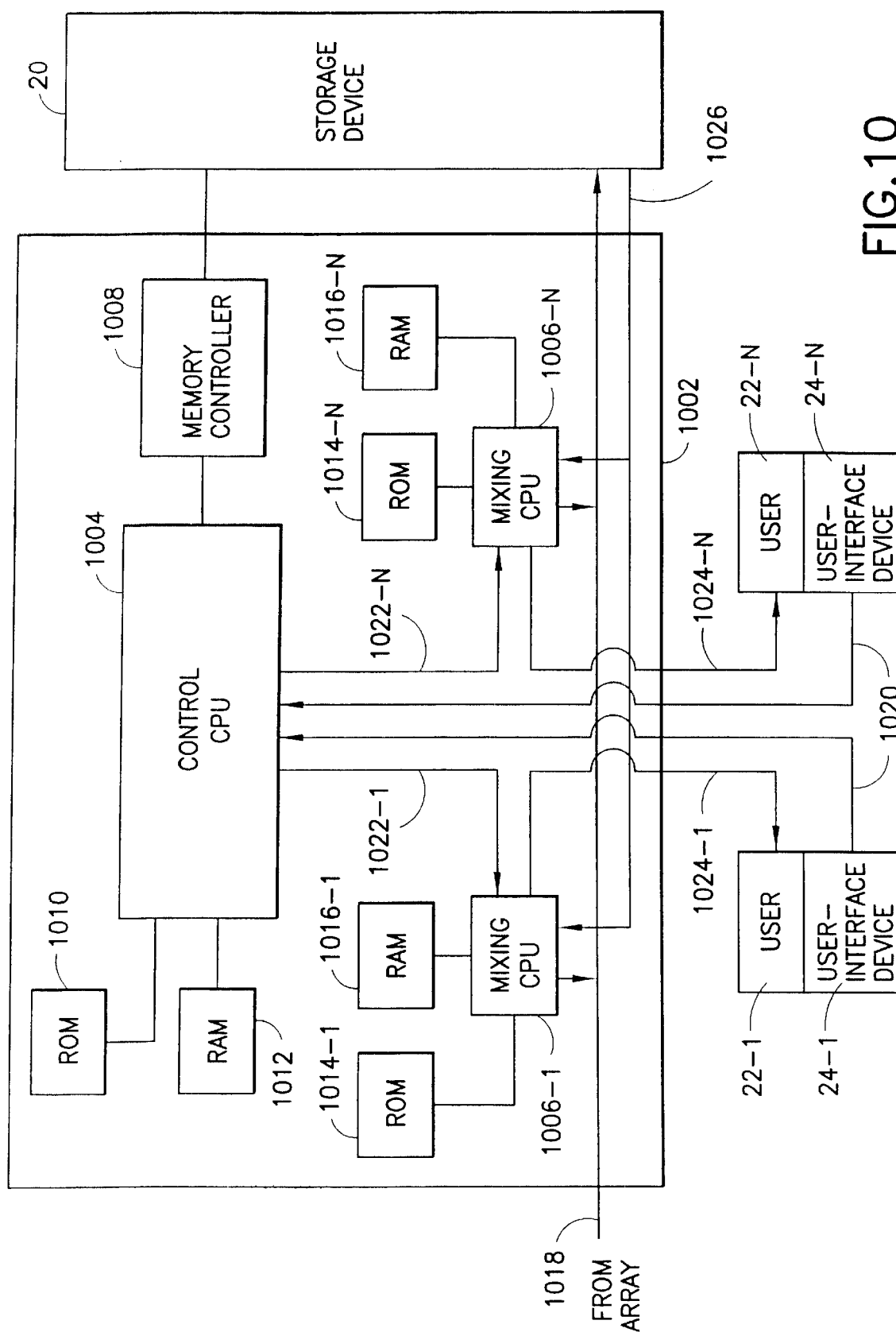


FIG. 10

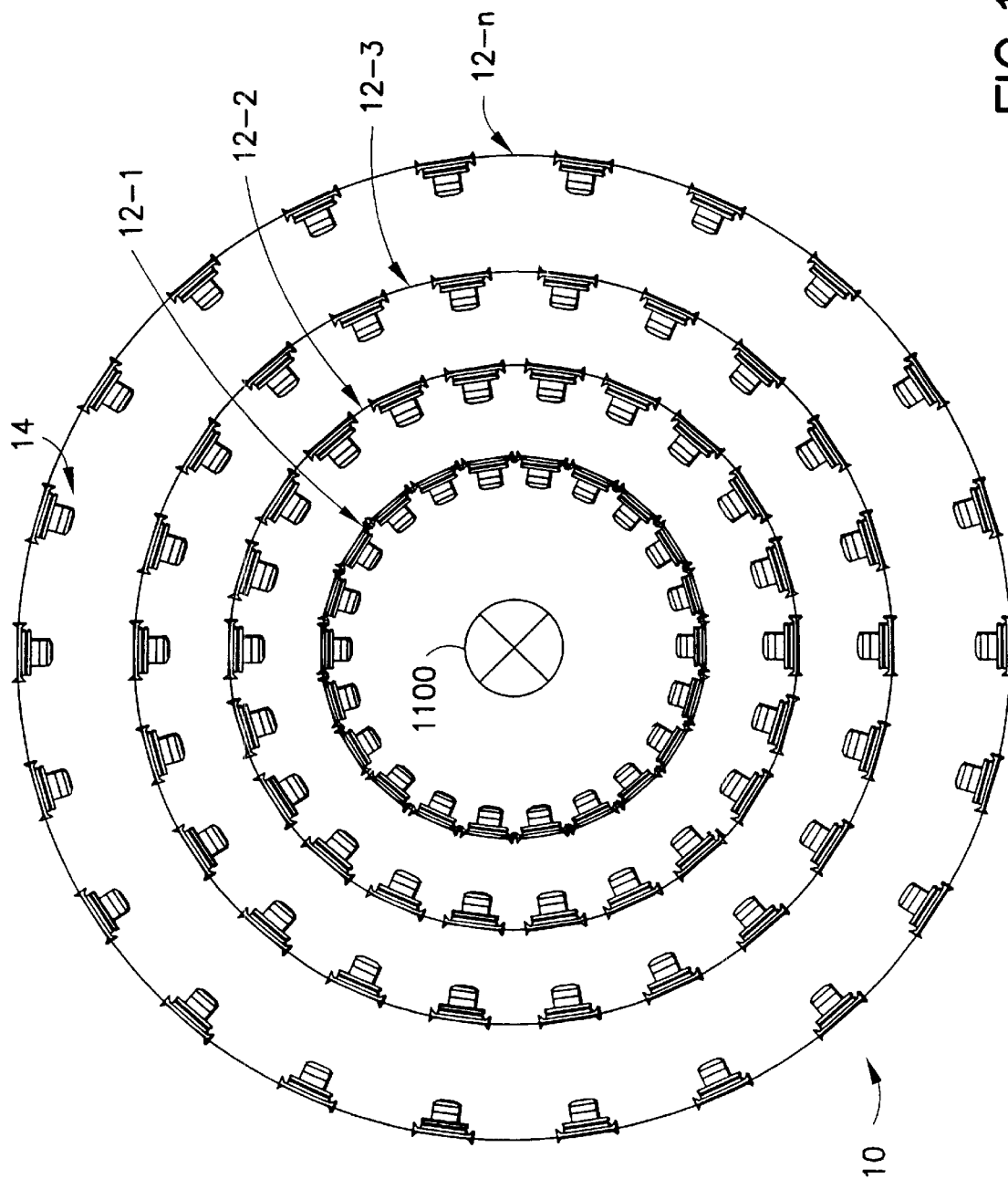


FIG. 11

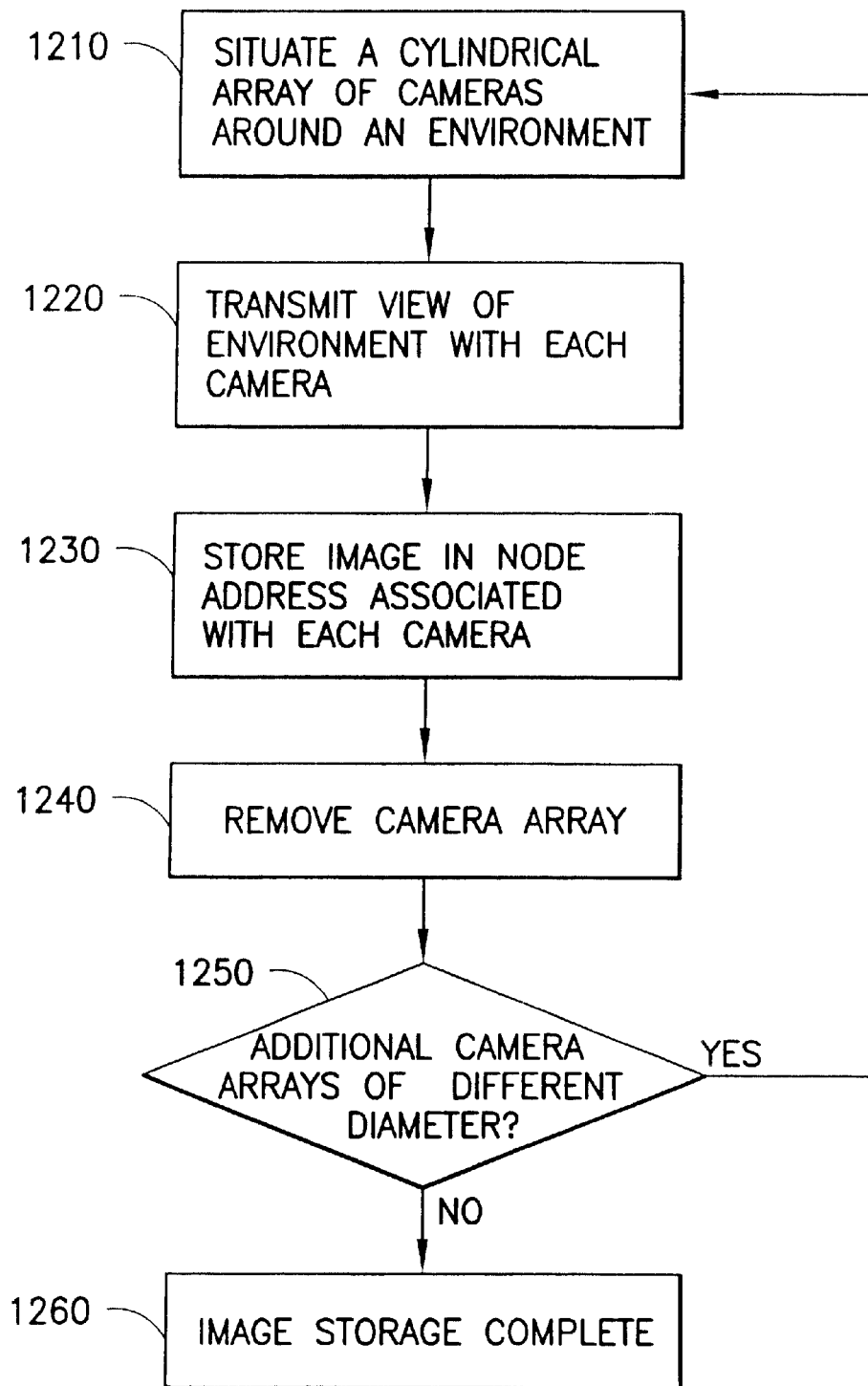


FIG. 12

NAVIGABLE TELEPRESENCE METHOD AND SYSTEM UTILIZING AN ARRAY OF CAMERAS

CROSS-REFERENCE TO RELATED APPLICATIONS

This application is a continuation-in-part of commonly assigned pending U.S. Pat. application Ser. No. 09/283,413 filed on Apr. 1, 1999 pending entitled A Navigable Telepresence Method and System Utilizing An Array of Cameras and claims priority under 35 U.S.C. §120 to such application which claims the benefit of U.S. Provisional Application Serial No. 60/080,413, filed on Apr. 2, 1998 pending both of which are herein incorporated by reference.

BACKGROUND OF THE INVENTION

1. Field Of The Invention

The present invention relates to a telepresence system and, more particularly, to a navigable camera array telepresence system and method of using same.

2. Description Of Related Art

In general, a need exists for the development of telepresence systems suitable for use with static venues, such as museums, and dynamic venues or events, such as music concerts. The viewing of such venues is limited by time, geographical location, and the viewer capacity of the venue. For example, potential visitors to a museum may be prevented from viewing an exhibit due to the limited hours the museum is open. Similarly, music concert producers must turn back fans due to the limited seating of an arena. In short, limited access to venues reduces the revenue generated.

In an attempt to increase the revenue stream from both static and dynamic venues, such venues have been recorded for broadcast or distribution. In some instances, dynamic venues are also broadcast live. While such broadcasting increases access to the venues, it involves considerable production effort. Typically, recorded broadcasts must be cut and edited, as views from multiple cameras are pieced together. These editorial and production efforts are costly.

In some instances, the broadcast resulting from these editorial and production efforts provides viewers with limited enjoyment. Specifically, the broadcast is typically based on filming the venue from a finite number of predetermined cameras. Thus, the broadcast contains limited viewing angles and perspectives of the venue. Moreover, the viewing angles and perspectives presented in the broadcast are those selected by a producer or director during the editorial and production process; there is no viewer autonomy. Furthermore, although the broadcast is often recorded for multiple viewings, the broadcast has limited content life because each viewing is identical to the first. Because each showing looks and sounds the same, viewers rarely come back for multiple viewings.

A viewer fortunate enough to attend a venue in person will encounter many of the same problems. For example, a museum-goer must remain behind the barricades, viewing exhibits from limited angles and perspectives. Similarly, concert-goers are often restricted to a particular seat or section in an arena. Even if a viewer were allowed free access to the entire arena to videotape the venue, such a recording would also have limited content life because each viewing would be the same as the first. Therefore, a need exists for a telepresence system that preferably provides user autonomy while resulting in recordings with enhanced content life at a reduced production cost.

Apparently, attempts have been made to develop telepresence systems to satisfy some of the foregoing needs. One telepresence system is described in U.S. Pat. No. 5,708,469 for Multiple View Telepresence Camera Systems Using A Wire Cage Which Surrounds A Polarity Of Multiple Cameras And Identifies The Fields Of View, issued Jan. 13, 1998. The system disclosed therein includes a plurality of cameras, wherein each camera has a field of view that is space-contiguous with and at a right angle to at least one other camera. In other words, it is preferable that the camera fields of view do not overlap each other. A user interface allows the user to jump between views. In order for the user's view to move through the venue or environment, a moving vehicle carries the cameras.

This system, however, has several drawbacks. For example, in order for a viewer's perspective to move through the venue, the moving vehicle must be actuated and controlled. In this regard, operation of the system is complicated. Furthermore, because the camera views are contiguous, typically at right angles, changing camera views results in a discontinuous image.

Other attempts at providing a telepresence system have taken the form of a 360 degree camera systems. One such system is described in U.S. Pat. No. 5,745,305 for Panoramic Viewing Apparatus, issued Apr. 28, 1998. The system described therein provides a 360 degree view of environment by arranging multiple cameras around a pyramid shaped reflective element. Each camera, all of which share a common virtual optical center, receives an image from a different side of the reflective pyramid. Other types of 360 degree camera systems employ a parabolic lens or a rotating camera.

Such 360 degree camera systems also suffer from drawbacks. In particular, such systems limit the user's view to 360 degrees from a given point perspective. In other words, 360 degree camera systems provide the user with a panoramic view from a single location. Only if the camera system was mounted on a moving vehicle could the user experience simulated movement through an environment.

U.S. Pat. No. 5,187,571 for Television System For Displaying Multiple Views of A Remote Location issued Feb. 16, 1993, describes a camera system similar to the 360 degree camera systems described above. The system described provides a user to select an arbitrary and continuously variable section of an aggregate field of view. Multiple cameras are aligned so that each camera's field of view merges contiguously with those of adjacent cameras thereby creating the aggregate field of view. The aggregate field of view may expand to cover 360 degrees. In order to create the aggregate field of view, the cameras' views must be contiguous. In order for the camera views to be contiguous, the cameras have to share a common point perspective, or vertex. Thus, like the previously described 360 degree camera systems, the system of U.S. Pat. No. 5,187,571 limits a user's view to a single point perspective, rather than allowing a user to experience movement in perspective through an environment.

Also, with regard to the system of U.S. Pat. No. 5,187,571, in order to achieve the contiguity between camera views, a relatively complex arrangement of mirrors is required. Additionally, each camera seemingly must also be placed in the same vertical plane.

Thus, a need still exists for an improved telepresence system that provides the ability to better simulate a viewer's actual presence in a venue, preferably in real time.

SUMMARY OF THE INVENTION

These and other needs are satisfied by the present invention. A telepresence system according to one embodiment of

the present invention includes an array of cameras, each of which has an associated view of an environment and an associated output representing the view. The system also includes a first user interface device having first user inputs associated with movement along a first path in the array. The system further includes a second user interface device having second user inputs associated with movement along a second path in the array. A processing element is coupled to the user interface devices. The processing element receives and interprets the first inputs and selects outputs of cameras in the first path. Similarly, the processing element receives and interprets the second inputs and selects outputs of cameras in the second path independently of the first inputs. Thus, a first user and a second user are able to navigate simultaneously and independently through the array. In another embodiment, the system may also mix the output by mosaicing or tweening the output images. In a further embodiment of the present invention the telepresence system distinguishes between permissible cameras in the array and impermissible cameras in the array. In yet another embodiment of the present invention the telepresence system allows a user to move forward or backward through the environment, which provides the user the opportunity to move forward or backward through the environment.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is an overall schematic of one embodiment of the present invention.

FIG. 2a is a perspective view of a camera and a camera rail section of the array according to one embodiment of the present invention.

FIGS. 2b—2d are side plan views of a camera and a camera rail according to one embodiment of the present invention.

FIG. 2e is a top plan view of a camera rail according to one embodiment of the present invention.

FIG. 3 is a perspective view of a portion of the camera array according to one embodiment of the present invention.

FIG. 4 is a perspective view of a portion of the camera array according to an alternate embodiment of the present invention.

FIG. 5 is a flowchart illustrating the general operation of the user interface according to one embodiment of the present invention.

FIG. 6 is a flowchart illustrating in detail a portion of the operation shown in FIG. 5.

FIG. 7a is a perspective view of a portion of one embodiment of the present invention illustrating the arrangement of the camera array relative to objects being viewed.

FIGS. 7b—7g illustrate views from the perspectives of selected cameras of the array in FIG. 7a.

FIG. 8 is a schematic view of an alternate embodiment of the present invention.

FIG. 9 is a schematic view of a server according to one embodiment of the present invention.

FIG. 10 is a schematic view of a server according to an alternate embodiment of the present invention.

FIG. 11 is a top plan view of an alternate embodiment of the present invention.

FIG. 12 is a flowchart illustrating in detail the image capture portion of the operation of the embodiment shown in FIG. 11.

DESCRIPTION OF PREFERRED EMBODIMENTS

1. General Description Of Preferred Embodiments

The present invention relates to a telepresence system that, in a preferred embodiments, uses modular, interlocking

arrays of microcameras. The cameras are on rails, with each rail holding a plurality of cameras. These cameras, each locked in a fixed relation to every adjacent camera on the array and dispersed dimensionally in a given environment, transmit image output to an associated storage node, thereby enabling remote viewers to navigate through such environment with the same moving light reflections and shadows) that characterize an actual in-environment transit.

In another preferred embodiment, the outputs of these microcameras are linked by tiny (less than half the width of a human hair) Vertical Cavity Surface Emitting Lasers (VCSELs) to optical fibers, fed through area net hubs, buffered on server arrays or server farms (either for recording or (instantaneous) relay) and sent to viewers at remote terminals, interactive wall screens, or mobile image appliances (like Virtual Retinal Displays). Each remote viewer, through an intuitive graphical user interface (GUI), can navigate effortlessly through the environment, enabling seamless movement through the event.

This involves a multiplexed, electronic switching process (invisible to the viewer) which moves the viewer's point perspective from camera to camera. Rather than relying, *per se*, on physically moving a microcamera through space, the system uses the multiplicity of positioned microcameras to move the viewer's perspective from microcamera node to adjacent microcamera node in a way that provides the viewer with a sequential visual and acoustical path throughout the extent of the array. This allows the viewer to fluidly track or dolly through a 3-dimensional remote environment, to move through an event and make autonomous real-time decisions about where to move and when to linger.

Instead of investing the viewer with the capacity to physically move a robotic camera, which would immediately limit the number of viewers that could simultaneously control their own course and navigate via storage nodes containing images of an environment associated with a pre-existing array of cameras. The user can move around the environment in any direction—clockwise or counterclockwise, up, down, closer to or further away from the environment, or some combination thereof. Moreover, image output mixing, such as mosaicing and tweening, effectuates seamless motion throughout the environment.

2. Detailed Description Of Preferred Embodiments

Certain embodiments of the present invention will now be described in greater detail with reference to the drawings. It is understood that the operation and functionality of many of the components of the embodiments described herein are known to one skilled in the art and, as such, the present description does not go into detail into such operative and functionality.

A telepresence system **100** according to the present invention is shown in FIG. 1. The telepresence system **100** generally includes an array **10** of cameras **14** coupled to a server **18**, which in turn is coupled to one or more users **22** each having a user interfaced/display device **24**. As will be understood to one skilled in the art, the operation and functionality of the embodiment described herein is provided, in part, by the server and user interface/display device. While the operation of these components is not described by way of particular code listings or logic diagrams, it is to be understood that one skilled in the art will be able to arrive at suitable implementations based on the functional and operational details provided herein. Furthermore, the scope of the present invention is not to be construed as limited to any particular code or logic implementation.

In the present embodiment, the camera array **10** is conceptualized as being in an X, Z coordinate system. This allows each camera to have an associated, unique node address comprising an X, and Z coordinate (X, Z). In the present embodiment, for example, a coordinate value corresponding to an axis of a particular camera represents the number of camera positions along that axis the particular camera is displaced from a reference camera. In the present embodiment, from the user's perspective the X axis runs left and right, and the Z axis runs down and up. Each camera **14** is identified by its X, Z coordinate. It is to be understood, however, that other methods of identifying cameras **14** can be used. For example, other coordinate systems, such as those noting angular displacement from a fixed reference point as well as coordinate systems that indicate relative displacement from the current camera node may be used. In another alternate embodiment, the array is three dimensional, located in an X, Y, Z coordinate system.

The array **10** comprises a plurality of rails **12**, each rail **12** including a series of cameras **14**. In the present preferred embodiment, the cameras **14** are microcameras. The output from the microcameras **14** are coupled to the server **18** by means of local area hubs **16**. The local area hubs **16** gather the outputs and, when necessary, amplify the outputs for transmission to the server **18**. In an alternate embodiment, the local area hubs **16** multiplex the outputs for transmission to the server **18**. Although the figure depicts the communication links **15** between the cameras **14** and the server **18** as being hardwired, it is to be understood that wireless links may be employed. Thus, it is within the scope of the present invention for the communication links **15** to take the form of fiber optics, cable, satellite, microwave transmission, internet, and the like.

Also coupled to the server **18** is an electronic storage device **20**. The server **18** transfers the outputs to the electronic storage device **20**. The electronic (mass) storage device **20**, in turn, transfers each camera's output onto a storage medium or means, such as CD-ROM, DVD, tape, platter, disk array, or the like. The output of each camera **14** is stored in a particular location on the storage medium associated with that camera **14** or is stored with an indication to which camera **14** each stored output corresponds. For example, the output of each camera **14** is stored in contiguous locations on a separate disk, tape, CD-ROM, or platter. As is known in the art, the camera output may be stored in a compressed format, such as JPEG, MPEG1, MPEG2, and the like. Having stored each output allows a user to later view the environment over and over again, each time moving through the array **10** in a new path, as described below. In some embodiments of the present invention, such as those providing only real-time viewing, no storage device is required.

As will be described in detail below, the server **18** receives output from the cameras **14** in the array. The server **18** processes these outputs for either storage in the electronic storage device **20**, transmission to the users **22** or both.

It is to be understood that although the server **18** is configured to provide the functionality of the system **100** in the present embodiment, it is to be understood that other processing elements may provide the functionality of the system **100**. For example, in alternate embodiments, the user interface device is a personal computer programmed to interpret the user input and transmit an indication of the desired current node address, buffer outputs from the array, and provide other of the described functions.

As shown, the system **100** can accommodate (but does not require) multiple users **22**. Each user **22** has associated

therewith a user interface device including a user display device (collectively **24**). For example, user **22-1** has an associated user interface device and a user display device in the form of a computer **24-1** having a monitor and a keyboard. User **22-2** has associated therewith an interactive wall screen **24-2** which serves as a user interface device and a user display device. The user interface device and the user display device of user **22-3** includes a mobile audio and image appliance **24-3**. A digital interactive TV **24-4** is the user interface device and user display device of user **22-4**. Similarly, user **22-5** has a voice recognition unit and monitor **24-5** as the user interface and display devices. It is to be understood that the foregoing user interface devices and user display devices are merely exemplary; for example, other interface devices include a mouse, touch screen, biofeedback devices, as well as those identified in U.S. Provisional Patent Application Serial No. 60/080,413 and the like.

As described in detail below, each user interface device **24** has associated therewith user inputs. These user inputs allow each user **22** to move or navigate independently through the array **10**. In other words, each user **22** enters inputs to generally select which camera outputs are transferred to the user display device. Preferably, each user display device includes a graphical representation of the array **10**. The graphical representation includes an indication of which camera in the array the output of which is being viewed. The user inputs allow each user to not only select particular cameras, but also to select relative movement or navigational paths through the array **10**.

As shown in FIG. 1, each user **22** may be coupled to the server **18** by an independent communication link. Furthermore, each communication link may employ different technology. For example, in alternate embodiments, the communication links include an internet link, a microwave signal link, a satellite link, a cable link, a fiber optic link, a wireless link, and the like.

It is to be understood that the array **10** provides several advantages. For example, because the array **10** employs a series of cameras **14**, no individual camera, or the entire array **10** for that matter, need be moved in order to obtain a seamless view of the environment. Instead, the user navigates through the array **10**, which is strategically placed through and around the physical environment to be viewed. Furthermore, because the cameras **14** of the array **10** are physically located at different points in the environment to be viewed, a user is able to view changes in perspective, a feature unavailable to a single camera that merely changes focal length.

Microcameras

Each camera **14** is preferably a microcamera. The microcameras—microlenses mounted on thumbnail-sized CMOS active pixel sensor (APS) microchips—are arranged in patterns that enable viewers to move radically, in straight lines, or in fluid combinations thereof. The cameras are produced in a mainstream manufacturing process, by several companies, including Photobit, Pasadena, Calif.; Sarnoff Corporation, Princeton, N.J.; and VLSI Vision, Ltd., Edinburgh, Scotland.

Structure of the Array

The structure of the array **10** will now be described in greater detail with reference to FIGS. 2a-2e. In general, the camera array **10** of the present embodiment comprises a series of modular rails **12** carrying microcameras **14**. The structure of the rails **12** and cameras **14** will now be

discussed in greater detail with reference to FIGS. 2a through 2d. Each camera 14 includes registration pins 34. In the preferred embodiment, the cameras 14 utilize VCSELs to transfer their outputs to the rail 12. It is to be understood that the present invention is not limited to any particular type of camera 14, however, or even to an array 10 consisting of only one type of camera 14.

Each rail 12 includes two sides, 12a, 12b, at least one of which 12b is hingeably connected to the base 12c of the rail 12. The base 12c includes docking ports 36 for receiving the registration pins 34 of the camera 14. When the camera 14 is seated on a rail 12 such that the registration pins 34 are fully engaged in the docking ports 36, the hinged side 12b of the rail 12 is moved against the base 32 of the camera 14, thereby securing the camera 14 to the rail 12.

Each rail 12 further includes a first end 38 and a second end 44. The first end 38 includes, in the present embodiment, two locking pins 40 and a protected transmission relay port 42 for transmitting the camera outputs. The second end 44 includes two guide holes 46 for receiving the locking pins 40, and a transmission receiving port 48. Thus, the first end 38 of one rail 12 is engagable with a second end 44 of another rail 12. Therefore, each rail 12 is modular and can be functionally connected to another rail to create the array 10.

Once the camera 14 is securely seated to the rail 12, the camera 14 is positioned such that the camera output may be transmitted via the VCSEL to the rail 12. Each rail 12 includes communication paths for transmitting the output from each camera 14.

Although the array 10 is shown having a particular configuration, it is to be understood that virtually any configuration of rails 12 and cameras 14 is within the scope of the present invention. For example, the array 10 may be a linear array of cameras 14, a 2-dimensional array of cameras 14, a 3-dimensional array of cameras 14, or any combination thereof. Furthermore, the array 10 need not be comprised solely of linear segments, but rather may include curvilinear sections.

The array 10 is supported by any of a number of support means. For example, the array 10 can be fixedly mounted to a wall or ceiling; the array 10 can be secured to a moveable frame that can be wheeled into position in the environment or supported from cables.

FIG. 3 illustrates an example of a portion of the array 10. As shown, the array 10 comprises five rows of rails 12a, through 12e. Each of these rails 12a–12e is directed towards a central plane, which substantially passes through the center row 12c. Consequently, for any object placed in the same plane as the middle row 12c, a user would be able to view the object essentially from the bottom, front, and top.

As noted above, the rails 12 of the array 10 need not have the same geometry. For example, some of the rails 12 may be straight while others may be curved. For example, FIG. 4 illustrates the camera alignment that results from utilizing curved rails. It should be noted that rails in FIG. 4 have been made transparent so that the arrangement of cameras 14 may be easily seen.

In an alternate embodiment, each rail is configured in a step-like fashion or an arc with each camera above (or below) and in front of a previous camera. In such an arrangement, the user has the option of moving forward through the environment.

It is to be understood that the spacing of the microcameras 14 depends on the particular application, including the objects being viewed, the focal length of the microcameras

14, and the speed of movement through the array 10. In one embodiment the distance between microcameras 14 can be approximated by analogy to a conventional movie reel recording projector. In general, the speed of movement of a projector through an environment divided by the frames per unit of time second results in a frame-distance ratio.

For example, as shown by the following equations, in some applications a frame is taken ever inch. A conventional movie projector records twenty-four frames per second. When such a or is moved through an environment at two feet per a frame is taken approximately every inch.

$$\begin{aligned} \frac{2 \text{ ft}}{\text{sec}} \div \frac{24 \text{ frames}}{\text{sec}} &= \\ \frac{2 \text{ ft}}{24 \text{ frames}} &= \frac{1 \text{ ft}}{12 \text{ frames}} = \frac{12 \text{ inches}}{12 \text{ frames}} = \frac{1 \text{ inch}}{1 \text{ frame}} = 1 \text{ frame per inch.} \end{aligned}$$

A frame of the projector is analogous to a camera 14 in the present invention. Thus, where one frame per inch results in a movie having a seamless view of the environment, so too does one camera 14 per inch. Thus, in one embodiment of the present invention the cameras 14 are spaced approximately one inch apart, thereby resulting in a seamless view of the environment.

Navigation Through the System

The general operation of the present embodiment will now be described with reference to FIG. 5 and continuing reference to FIG. 1. As shown in step 110, the user is presented with a predetermined starting view of the environment corresponding to a starting camera. It is to be understood that the operation of the system is controlled, in part, by software residing in the server. As noted above, the system associates each camera in the array with a coordinate. Thus, the system is able to note the coordinates of the starting camera node. The camera output and, thus the corresponding view, changes only upon receiving a user input.

When the user determines that they want to move or navigate through the array, the user enters a user input through the user interface device 24. As described below, the user inputs of the present embodiment generally include moving to the right, to the left, up, or down in the array. Additionally, a user may jump to a particular camera in the array. In alternate embodiments, a subset of these or other inputs, such as forward, backward, diagonal, over, and under, are used. The user interface device, in turn, transmits the user input to the server in step 120.

Next, the server receives the user input in step 130 and proceeds to decode the input. In the present embodiment, decoding the input generally involves determining whether the user wishes to move to the right, to the left, up, or down in the array.

On the other hand, if the received user input does not correspond to backward, then The server 18 proceeds to determine whether the input corresponds to moving to the user's right in the array 10. This determination is shown in step 140. If the received user input does correspond to moving to the right, the current node address is incremented along the X axis in step 150 to obtain an updated address.

If the received user input does not correspond to moving to the right in the array, the server 18 then determines whether the input corresponds to moving to the user's left in the array 10 in step 160. Upon determining that the input does correspond to moving to the left, the server 18 then

decrements the current node address along the X axis to arrive at the updated address. This is shown in step 170.

If the received user input does not correspond to either moving to the right or to the left, the server 18 then determines whether the input corresponds to moving up in the array. This determination is made in step 180. If the user input corresponds to moving up, in step 190, the server 18 increments the current node address along the Z axis, thereby obtaining an updated address.

Next, the server 18 determines whether the received user input corresponds to moving down in the array 10. This determination is made in step 200. If the input does correspond to moving down in the array 10, in step 210 the server 18 decrements the current node address along the Z axis.

Lastly, in step 220 the server 18 determines whether the received user input corresponds to jumping or changing the view to a particular camera 14. As indicated in FIG. 5, if the input corresponds to jumping to a particular camera 14, the server 18 changes the current node address to reflect the desired camera position. Updating the node address is shown as step 230. In an alternate embodiment, the input corresponds to jumping to a particular position in the array 10, not identified by the user as being a particular camera but by some reference to the venue, such as stage right.

It is to be understood that the server 18 may decode the received user inputs in any of a number of ways, including in any order. For example, in an alternate embodiment the server 18 first determines whether the user input corresponds to up or down. In another alternate, preferred embodiment, user navigation includes moving forward, backward, to the left and right, and up and down through a three dimensional array.

If the received user input does not correspond to any of the recognized inputs, namely to the right, to the left, up, down, or jumping to a particular position in the array 10 then in step 240, the server 18 causes a message signal to be transmitted to the user display device 24, causing a message to be displayed to the user 22 that the received input was not understood. Operation of the system 100 then continues with step 120, and the server 18 awaits receipt of the next user input.

After adjusting the current node address, either by incrementing or decrementing the node address along an axis or by jumping to a particular node address, the server 18 proceeds in step 250 to adjust the user's view. Once the view is adjusted, operation of the system 100 continues again with step 120 as the server 18 awaits receipt of the next user input.

In an alternate embodiment, the server 18 continues to update the node address and adjust the view based on the received user input. For example, if the user input corresponded to "moving to the right", then operation of the system 100 would continuously loop through steps 140, 150, and 250, checking for a different input. When the different input is received, the server 18 continuously updates the view accordingly.

It is to be understood that the foregoing user inputs, namely, to the right, to the left, up, and down, are merely general descriptions of movement through the array. Although the present invention is not so limited, in the present preferred embodiment, movement in each of these general directions is further defined based upon the user input.

Accordingly, FIG. 6 is a more detailed diagram of the operation of the system according to steps 140, 150, and 250 of FIG. 5. Moreover, it is to be understood that while FIG.

6 describes more detailed movement one direction i.e., to the right, the same detailed movement can be applied in any other direction. As illustrated, the determination of whether the user input corresponds to moving to the right actually involves several determinations. As described in detail below, these determinations include moving to the right through the array 10 at different speeds, moving to the right into a composited additional source output at different speeds, and having the user input overridden by the system 100.

The present invention allows a user 22 to navigate through the array 10 at the different speeds. Depending on the speed (i.e. number of camera nodes transversed per unit of time) indicated by the user's input, such as movement of a pointing device (or other interface device), the server 18 will apply an algorithm that controls the transition between camera outputs either at critical speed (n nodes/per unit of time), under critical speed ($n-1$ nodes/per unit of time), or over critical speed ($n+1$ nodes/per unit of time).

It is to be understood that speed of movement through the array 10 can alternatively be expressed as the time to switch from one camera 14 to another camera 14.

Specifically, as shown in step 140a, the server 18 makes the determination whether the user input corresponds to moving to the right at a critical speed. The critical speed is preferably a predetermined speed of movement through the array 10 set by the system operator or designer depending on the anticipated environment being viewed. Further, the critical speed depends upon various other factors, such as focal length, distance between cameras, distance between the cameras and the viewed object, and the like. The speed of movement through the array 10 is controlled by the number of cameras 14 traversed in a given time period. Thus, the movement through the array 10 at critical speed corresponds to traversing some number, " n ", camera nodes per millisecond, or taking some amount of time, " s ", to switch from one camera 14 to another. It is to be understood that in the same embodiment the critical speed of moving through the array 10 in one dimension need not equal the critical speed of moving through the array in another dimension. Consequently, the server 18 increments the current node address along the X axis at n nodes per millisecond.

In the present preferred embodiment the user traverses twenty-four cameras 14 per second. As discussed above, a movie projector records twenty-four frames per second. Analogizing between the movie projector and the present invention, at critical the user traverses (and the server 18 switches between) approximately twenty-four cameras 14 per second, or a camera 14 approximately every 0.04167 seconds.

As shown in FIG. 6, the user 22 may advance not only at critical speed, but also at over the critical speed, as shown in step 140b, or at under the critical speed, as shown in step 140c. Where the user input "I" indicates movement through the array 10 at over the critical speed, the server 18 increments the current node address along the X axis by a unit of greater than n , for example, at $n+2$ nodes per millisecond. The step of incrementing the current node address at $n+1$ nodes per millisecond along the X axis is shown in step 150b. Where the user input "I" indicates movement through the array 10 at under the critical speed, the server 18 proceeds to increment the current node address at a variable less than n , for example, at $n-1$ nodes per millisecond. This operation is shown as step 150c.

Scaleable Arrays

The shape of the array 10 can also be electronically scaled and the system 100 designed with a "center of gravity" that

will ease a user's image path back to a "starting" or "critical position" node or ring of nodes, either when the user 22 releases control or when the system 100 is programmed to override the user's autonomy; that is to say, the active perimeter or geometry of the array 10 can be pre-configured to change at specified times or intervals in order to corral or focus attention in a situation that requires dramatic shaping. The system operator can, by real-time manipulation or via a pre-configured electronic proxy sequentially activate or deactivate designated portions of the camera array 10. This is of particular importance in maintaining authorship and dramatic pacing in theatrical or entertainment venues, and also for implementing controls over how much freedom a user 22 will have to navigate through the array 10.

In the present embodiment, the system 100 can be programmed such that certain portions of the array 10 are unavailable to the user 22 at specified times or intervals. Thus, continuing with step 140d of FIG. 6, the server 18 makes the determination whether the user input corresponds to movement to the right through the array but is subject to a navigation control algorithm. The navigation control algorithm causes the server 18 to determine, based upon navigation control factors, whether the user's desired movement is permissible.

More specifically, the navigation control algorithm, which is programmed in the server 18, determines whether the desired movement would cause the current node address to fall outside the permissible range of node coordinates. In the present embodiment, the permissible range of node coordinates is predetermined and depends upon the time of day, as noted by the server 18. Thus, in the present embodiment, the navigation control factors include time. As will be appreciated by those skilled in the art, permissible camera nodes and control factors can be correlated in a table stored in memory.

In an alternate embodiment, the navigation control factors include time as measured from the beginning of a performance being viewed, also as noted by the server. In such an embodiment, the system operator can dictate from where in the array a user will view certain scenes. In another alternate embodiment, the navigation control factor is speed of movement through the array. For example, the faster a user 22 moves or navigates through the array, the wider the turns must be. In other alternate embodiments, the permissible range of node coordinates is not predetermined. In one embodiment, the navigation control factors and, therefore, the permissible range, is dynamically controlled by the system operator who communicates with the server via an input device.

Having determined that the user input is subject to the navigation control algorithm, the server 18 further proceeds, in step 150d, to increment the current node address along a predetermined path. By incrementing the current node address along a predetermined path, the system operator is able to corral or focus the attention of the user 22 to the particular view of the permissible cameras 14, thereby maintaining authorship and dramatic pacing in theatrical and entertainment venues.

In an alternate embodiment where the user input is subject to a navigation control algorithm, the server 18 does not move the user along a predetermined path. Instead, the server 18 merely awaits a permissible user input and holds the view at the current node. Only when the server 18 receives a user input resulting in a permissible node coordinate will the server 18 adjust the user's view.

Additional Source Output

In addition to moving through the array 10, the user 22 may, at predetermined locations in the array 10, choose to

leave the real world environment being viewed. More specifically, additional source outputs, such as computer graphic imagery, virtual world imagery, applets, film clips, and other artificial and real camera outputs, are made available to the user 22. In one embodiment, the additional source output is composited with the view of the real environment. In an alternate embodiment, the user's view transfers completely from the real environment to that offered by the additional source output.

More specifically, the additional source output is stored (preferably in digital form) in the electronic storage device 20. Upon the user 22 inputting a desire to view the additional source output, the server 18 transmits the additional source output to the user interface/display device 24. The present embodiment, the server 18 simply transmits the additional source output to the user display device 24. In an alternate embodiment, the server 18 first composites the additional source output with the camera output and then transmits the composited signal to the user interface/display device 24.

As shown in step 140e, the server 18 makes the determination whether the user input corresponds to moving in the array into the source output. If the user 22 decides to move into the additional source output, the server 18 adjusts the view by substituting the additional source output for the updated camera output identified in either of steps 150a-d.

Once the current node address is updated in either of steps 150a-d, the server 18 proceeds to adjust the user's view in step 250. When adjusting the view, the server 18 "mixes" the existing or current camera output being displayed with the output of the camera 14 identified by the updated camera node address. Mixing the outputs is achieved differently in alternate embodiments of the invention. In the present embodiment, mixing the outputs involves electronically switching at a particular speed from the existing camera output to the output of the camera 14 having the new current node address.

It is to be understood that in this and other preferred embodiments disclosed herein, the camera outputs are synchronized. As is well known in the art, a synchronizing signal from a "sync generator" is supplied to the cameras. The sync generator may take the form of those used in video editing and may comprise, in alternate embodiments, part of the server, the hub, and/or a separate component coupled to the array.

As described above, at critical speed, the server 18 switches camera outputs approximately at a rate of 24 per second, or one every 0.04167 seconds. If the user 22 is moving through the array 10 at under the critical speed, the outputs of the intermediate cameras 14 are each displayed for a relatively longer duration than if the user is moving at the critical speed. Similarly, each output is displayed for a relatively shorter duration when a user navigates at over the critical speed. In other words, the server 18 adjusts the switching speed based on the speed of the movement through the array 10.

Of course, it is to be understood that in a simplified embodiment of the present invention, the user may navigate at only the critical speed.

In another alternate embodiment, mixing the outputs is achieved by compositing the existing or current output and the updated camera node output. In yet another embodiment, mixing involves dissolving the existing view into the new view. In still another alternate embodiment, mixing the outputs includes adjusting the frame refresh rate of the user display device. Additionally, based on speed of movement through the array, the server may add motion blur to convey the realistic sense of speed.

In yet another alternate embodiment, the server causes a black screen to be viewed instantaneously between camera views. Such an embodiment is analogous to blank film between frames in a movie reel. Furthermore, although not always advantageous, such black screens reduce the physiologic “carrying over” of one view into a subsequent view.

It is to be understood that the user inputs corresponding to movements through the array at different speeds may include either different keystrokes on a keypad, different positions of a joystick, positioning a joystick in a given position for a predetermined length of time, and the like. Similarly, the decision to move into an additional source output may be indicated by a particular keystroke, joystick movement, or the like.

In another embodiment, mixing may be accomplished by “mosaicing” the outputs of the intermediate cameras **14**. U.S. Pat. No. 5,649,032 entitled System For Automatically Aligning Images To Form A Mosaic Image to Peter J. Burt et al. discloses a system and method for generating a mosaic from a plurality of images and is hereby incorporated by reference. The server **18** automatically aligns one camera output to another camera output, a camera output to another mosaic (generated from previously occurring camera output) such that the output can be added to the mosaic, or an existing mosaic to a camera output.

Once the mosaic alignment is complete, the present embodiment utilizes a mosaic composition process to construct (or update) a mosaic. The mosaic composition comprises a selection process and a combination process. The selection process automatically selects outputs for incorporation into the mosaic and may include masking and cropping functions to select the region of interest in a mosaic. Once the selection process selects which output(s) are to be included in the mosaic, the combination process combines the various outputs to form the mosaic. The combination process applies various output processing techniques, such as merging, fusing, filtering, output enhancement, and the like, to achieve a seamless combination of the outputs. The resulting mosaic is a smooth view that combines the constituent outputs such that temporal and spatial information redundancy are minimized in the mosaic. In one embodiment of the present invention, the mosaic may be formed as the user moves through the system (on the fly) and the output image displayed close to real time. In another embodiment, the system may form the mosaic from a predetermined number of outputs or during a predetermined time interval, and then display the images pursuant to the user’s navigation through the environment.

In yet another embodiment, the server **18** enables the output to be mixed by a “tweening” process. One example of the tweening process is disclosed in U.S. Pat. No. 5,259,040 entitled Method For Determining Sensor Motion And Scene Structure And Image Processing System Therefor to Keith J. Hanna, herein incorporated by reference. Tweening enables the server **18** to process the structure of a view from two or more camera outputs of the view.

Applying the Hanna patent to the telepresence method/system herein, tweening is now described. The server monitors the movement among the intermediate cameras **14** through a scene using local scene characteristics such as brightness derivatives of a pair of camera outputs. A global camera output movement constraint is combined with a local scene characteristic constancy constraint to relate local surface structures with the global camera output movement model and local scene characteristics. The method for determining a model for global camera output movement through

a scene and scene structure model of the scene from two or more outputs of the scene at a given image resolution comprises the following steps:

- (a) setting initial estimates of local scene models and a global camera output movement model;
- (b) determining a new value of one of the models by minimizing the difference between the measured error in the outputs and the error predicted by the model;
- (c) resetting the initial estimates of the local scene models and the image sensor motion model using the new value of one of the models determined in step (b);
- (d) determining a new value of the second of the models using the estimates of the models determined in step (b) by minimizing the difference between the measured error in the outputs and the error predicted by the model;
- (e) warping one of the outputs towards the other output using the current estimates of the models at the given image resolution; and
- (f) repeating steps (b), (c), (d) and (e) until the differences between the new values of the models and the values determined in the previous iteration are less than a certain value or until a fixed number of iterations have occurred.

It should be noted that where the Hanna patent effectuates the tweening process by detecting the motion of an image sensor (e.g., a video camera), an embodiment of the present invention monitors the user movement among live cameras or storage nodes.

In an alternate embodiment, although not always necessary, to ensure a seamless progression of views, the server **18** also transmits to the user display device **24** outputs from some or all of the intermediate cameras, namely those located between the current camera node and the updated camera node. Such an embodiment will now be described with reference to FIGS. **7a–7g**. Specifically, FIG. **7a** illustrates a curvilinear portion of an array **10** that extends along the X axis or to the left and right from the user’s perspective. Thus, the coordinates that the server **18** associates with the cameras **14** differ only in the X coordinate. More specifically, for purposes of the present example, the cameras **14** can be considered sequentially numbered, starting with the left-most camera **14** being the first, i.e., number “1”. The X coordinate of each camera **14** is equal to the camera’s position in the array. For illustrative purposes, particular cameras will be designate **14-X**, where X equals the camera’s position in the array **10** and, thus, its associated X coordinate.

In general, FIGS. **7a–7g** illustrate possible user movement through the array **10**. The environment to be viewed includes three objects **602**, **604**, **606**, the first and second of which include numbered surfaces. As will be apparent, these numbered surface allow a better appreciation of the change in user perspective.

In FIG. **7a**, six cameras **14-2**, **14-7**, **14-11**, **14-14**, **14-20**, **14-23** of the array **10** are specifically identified. The boundaries of each camera’s view is identified by the pair of lines **14-2a**, **14-7a**, **14-11a**, **14-14a**, **14-20a**, **14-23a**, radiating from each identified camera **14-2**, **14-7**, **14-11**, **14-14**, **14-20**, **14-23**, respectively. As described below, in the present example the user **22** navigates through the array **10** along the X axis such that the images or views of the environment are those corresponding to the identified cameras **14-2**, **14-7**, **14-11**, **14-14**, **14-20**, **14-23**.

The present example provides the user **22** with the starting view from camera **14-2**. This view is illustrated in FIG. **7b**.

15

The user 22, desiring to have a better view of the object 702, pushes the "7" key on the keyboard. This user input is transmitted to and interpreted by the server 18.

Because the server 18 has been programmed to recognize the "7" key as corresponding to moving or jumping through the array to camera 14-7. The server 18 changes the X coordinate of the current camera node address to 7, selects the output of camera 14-7, and adjusts the view or image sent to the user 22. Adjusting the view, as discussed above, involves mixing the outputs of the current and updated camera nodes. Mixing the outputs, in turn, involves switching intermediate camera outputs into the view to achieve the seamless progression of the discrete views of cameras 14-2 through 14-7, which gives the user 22 the look and feel of moving around the viewed object. The user 22 now has another view of the first object 702. The view from camera 14-7 is shown in FIG. 7c. As noted above, if the jump in camera nodes is greater than a predetermined limit, the server 18 would omit some or all of the intermediate outputs.

Pressing the "right arrow" key on the keyboard, the user 22 indicates to the system 100 a desire to navigate to the right at critical speed. The server 18 receives and interprets this user input as indicating such and increments the current camera node address by n=4. Consequently, the updated camera node address is 14-11. The server 18 causes the mixing of the output of camera 14-11 with that of camera 14-7. Again, this includes switching into the view the outputs of the intermediate cameras (i.e., 14-8, 14-9, and 14-10) to give the user 22 the look and feel of navigating around the viewed object. The user 22 is thus presented with the view from camera 14-11, as shown in FIG. 7d.

Still interested in the first object 702, the user 22 enters a user input, for example, "alt-right arrow," indicating a desire to move to the right at less than critical speed. Accordingly, the server 18 increments the updated camera node address by n-1 nodes, namely 3 in the present example, to camera 14-14. The outputs from cameras 14-11 and 14-14 are mixed, and the user 22 is presented with a seamless view associated with cameras 14-11 through 14-14. FIG. 7e illustrates the resulting view of camera 14-14.

With little to see immediately after the first object 702, the user 22 enters a user input such as "shift-right arrow," indicating a desire to move quickly through the array 10, i.e., at over the critical speed. The server 18 interprets the user input and increments the current node address by n+2, or 6 in the present example. The updated node address thus corresponds to camera 14-20. The server 18 mixes the outputs of cameras 14-14 and 14-20, which includes switching into the view the outputs of the intermediate cameras 14-15 through 14-19. The resulting view of camera 14-20 is displayed to the user 22. As shown in FIG. 7f, the user 22 now views the second object 704.

Becoming interested in the third object 704, the user 22 desires to move slowly through the array 10. Accordingly, the user 22 enters "alt-right arrow" to indicate moving to the right at below critical speed. Once the server 18 interprets the received user input, it updates the current camera node address along the X axis by 3 to camera 14-23. The server 18 then mixes the outputs of camera 14-20 and 14-23, thereby providing the user 22 with a seamless progression of views through camera 14-23. The resulting view 14-23a is illustrated in FIG. 7g.

Other Data Devices

It is to be understood that devices other than cameras may be interspersed in the array. These other devices, such as motion sensors and microphones, provide data to the server

16

(s) for processing. For example, in alternate embodiments output from motion sensors or microphones are fed to the server(s) and used to scale the array. More specifically, permissible camera nodes (as defined in a table stored in memory) are those near the sensor or microphone having a desired output e.g., where there is motion or sound. As such, navigation control factors include output from other such devices. Alternatively, the output from the sensors or microphones are provided to the user.

An alternate embodiment in which the array of cameras includes multiple microphones interspersed among the viewed environment and the cameras will now be described with reference to FIG. 8. The system 800 generally includes an array of cameras 802 coupled to a server 804, which, in turn, is coupled to one or more user interface and display devices 806 and an electronic storage device 808. A hub 810 collects and transfers the outputs from the array 802 to the server 804. More specifically, the array 802 comprises modular rails 812 that are interconnected. Each rail 812 carries multiple microcameras 814 and a microphone 816 centrally located at rail 812. Additionally, the system 800 includes microphones 818 that are physically separate from the array 802. The outputs of both the cameras 814 and microphones 816, 818 are coupled to the server 804 for processing.

In general, operation of the system 800 proceeds as described with respect to system 100 of FIGS. 1-2d and 5-6. Beyond the operation of the previously described system 100, however, the server 804 receives the sound output from the microphones 816, 818 and, as with the camera output, selectively transmits sound output to the user. As the server 804 updates the current camera node address and changes the user's view, it also changes the sound output transmitted to the user. In the present embodiment, the server 804 has stored in memory an associated range of camera nodes with a given microphone, namely the cameras 814 on each rail 810 are associated with the microphone 816 on that particular rail 810. In the event a user attempts to navigate beyond the end of the array 802, the server 804 determines the camera navigation is impermissible and instead updates the microphone node output to that of the microphone 818 adjacent to the array 802.

In an alternate embodiment, the server 804 might include a database in which camera nodes in a particular area are associated with a given microphones. For example, a rectangle defined by the (X, Y, Z) coordinates (0, 0, 0), (10, 0, 0), (10, 5, 0), (0, 5, 0), (0, 0, 5), (10, 0, 5), (10, 5, 5) and (0, 5, 5) are associated with a given microphone. It is to be understood that selecting one of the series of microphones based on the user's position (or view) in the array provides the user with a sound perspective of the environment that coincides with the visual perspective.

It is to be understood that the server of the embodiments discussed above may take any of a number of known configurations. Two examples of server configurations suitable for use with the present invention will be described with reference to FIGS. 9 and 10. Turning first to FIG. 9, the server 902, electronic storage device 20, array 10, users (1, 2, 3, . . . N) 22-1-22-N, and associated user interface/display devices 24-1-24-N are shown therein.

The server 902 includes, among other components, a processing means in the form of one or more central processing units (CPU) 904 coupled to associated read only memory (ROM) 906 and a random access memory (RAM) 908. In general, ROM 906 is for storing the program that dictates the operation of the server 902, and the RAM 908

is for storing variables and values used by the CPU 904 during operation. Also coupled to the CPU 904 are the user interface/display devices 24. It is to be understood that the CPU may, in alternate embodiments, comprise several processing units, each performing a discrete function.

Coupled to both the CPU 904 and the electronic storage device 20 is a memory controller 910. The memory controller 910, under direction of the CPU 904, controls accesses (reads and writes) to the storage device 20. Although the memory controller 910 is shown as part of the server 902, it is to be understood that it may reside in the storage device 20.

During operation, the CPU 904 receives camera outputs from the array 10 via bus 912. As described above, the CPU 904 mixes the camera outputs for display on the user interface/display device 24. Which outputs are mixed depends on the view selected by each user 22. Specifically, each user interface/display devices 24 transmits across bus 914 the user inputs that define the view to be displayed. Once the CPU 904 mixes the appropriate outputs, it transmits the resulting output to the user interface/display device 24 via bus 916. As shown, in the present embodiment, each user 22 is independently coupled to the server 902.

The bus 912 also carries the camera outputs to the storage device 20 for storage. When storing the camera outputs, the CPU 904 directs the memory controller 910 to store the output of each camera 14 in a particular location of memory in the storage device 20.

When the image to be displayed has previously been stored in the storage device 20, the CPU 904 causes the memory controller 910 to access the storage device 20 to retrieve the appropriate camera output. The output is thus transmitted to the CPU 904 via bus 918 where it is mixed. Bus 918 also carries additional source output to the CPU 904 for transmission to the users 22. As with outputs received directly from the array 10, the CPU 904 mixes these outputs and transmits the appropriate view to the user interface/display device 24.

FIG. 10 shows a server configuration according to an alternate embodiment of the present invention. As shown therein, the server 1002 generally comprises a control central processing unit (CPU) 1004, a mixing CPU 1006 associated with each user 22, and a memory controller 1008. The control CPU 1004 has associated ROM 1010 and RAM 1012. Similarly, each mixing CPU 1006 has associated ROM 1014 and RAM 1016.

To achieve the functionality described above, the camera outputs from the array 10 are coupled to each of the mixing CPUs 1 through N 1006-1, 1006-N via bus 1018. During operation, each user 22 enters inputs in the interface/display device 24 for transmission (via bus 1020) to the control CPU 1004. The control CPU 1004 interprets the inputs and, via buses 1022-1, 1022-N, transmits control signals to the mixing CPUs 1006-1, 1006-N instructing them which camera outputs received on bus 1018 to mix. As the name implies, the mixing CPUs 1006-1, 1006-N mix the outputs in order to generate the appropriate view and transmit the resulting view via buses 1024-1, 1024-N to the user interface/display devices 24-1, 24-N.

In an alternate related embodiment, each mixing CPU 1006 multiplexes outputs to more than one user 22. Indications of which outputs are to be mixed and transmitted to each user 22 comes from the control CPU 1004.

The bus 1018 couples the camera outputs not only to the mixing CPUs 1006-1, 1006-N, but also to the storage device 20. Under control of the memory controller 1008, which in

turn is controlled by the control CPU 1004, the storage device 20 stores the camera outputs in known storage locations. Where user inputs to the control CPU 1004 indicate a users' 22 desire to view stored images, the control CPU 1004 causes the memory controller 1008 to retrieve the appropriate images from the storage device 20. Such images are retrieved into the mixing CPUs 1006 via bus 1026. Additional source output is also retrieved to the mixing CPUs 1006-1, 1006-N via bus 1026. The control CPU 1004 also passes control signals to the mixing CPUs 1006-1, 1006-N to indicate which outputs are to be mixed and displayed.

Stereoscopic Views

It is to be understood that it is within the scope of the present invention to employ stereoscopic views of the environment. To achieve the stereoscopic view, the system retrieves from the array (or the electronic storage device) and simultaneously transmits to the user at least portions of outputs from two cameras. The server processing element mixes these camera outputs to achieve a stereoscopic output. Each view provided to the user is based on such a stereoscopic output. In one stereoscopic embodiment, the outputs from two adjacent cameras in the array are used to produce one stereoscopic view. Using the notation of FIGS. 7a-7g, one view is the stereoscopic view from cameras 14-1 and 14-2. The next view is based on the stereoscopic output of cameras 14-2 and 14-3 or two other cameras. Thus, in such an embodiment, the user is provided the added feature of a stereoscopic seamless view of the environment.

Multiple Users

As described above, the present invention allows multiple users to simultaneously navigate through the array independently of each other. To accommodate multiple users, the systems described above distinguish between inputs from the multiple users and selects a separate camera output appropriate to each user's inputs. In one such embodiment, the server tracks the current camera node address associated with each user by storing each node address in a particular memory location associate with that user. Similarly, each user's input is differentiated and identified as being associated with the particular memory location with the use of message tags appended to the user inputs by the corresponding user interface device.

In an alternate embodiment, two or more users may choose to be linked, thereby moving in tandem and having the same view of the environment. In such an embodiment, each includes identifying another user by his/her code to serve as a "guide". In operation, the server provides the outputs and views selected by the guide user to both the guide and the other user selecting the guide. Another user input causes the server to unlink the users, thereby allowing each user to control his/her own movement through the array.

Multiple Arrays

In certain applications, a user may also wish to navigate forward and backward through the environment, thereby moving closer to or further away from an object. Although it is within the scope of the present invention to use cameras with zoom capability, simply zooming towards an object does not change the user's image point perspective. One such embodiment in which users can move dimensionally forward and backward through the environment with a changing image point perspective will now be described

with respect to FIG. 11 and continuing reference to FIG. 1. As will be understood by those skilled in the art, the arrays described with reference to FIG. 11 may be used with any server, storage device and user terminals described herein.

FIG. 11 illustrates a top plan view of another embodiment enabling the user to move left, right, up, down, forward or backwards through the environment. A plurality of cylindrical arrays (121-1-121-n) of differing diameters comprising a series of cameras 14 may be situated around an environment comprising one or more objects 1200, one cylindrical array at a time. Cameras 14 situated around the object(s) 1100 are positioned along an X and Z coordinate system. Accordingly, an array 12 may comprise a plurality of rings of the same circumference positioned at different positions (heights) throughout the z-axis to form a cylinder of cameras 14 around the object(s) 1100. This also allows each camera in each array 12 to have an associated, unique storage node address comprising an X and Z coordinate—i.e., arrayl(X, Z). In the present embodiment, for example, a coordinate value corresponding to an axis of a particular camera represents the number of camera positions along that axis the particular camera is displaced from a reference camera. In the present embodiment, from the user's perspective, the X axis runs around the perimeter of an array 12, and the Z axis runs down and up. Each storage node is associated with a camera view identified by its X, Z coordinate.

As described above, the outputs of the cameras 14 are coupled to one or more servers for gathering and transmitting the outputs to the server 18.

In one embodiment, because the environment is static, each camera requires only one storage location. The camera output may be stored in a logical arrangement, such as a matrix of n arrays, wherein each array has a plurality of (X, Z) coordinates. In one embodiment, the node addresses may comprise of a specific coordinate within an array—i.e., Array₁(X_n, Z_n), Array₂(X_n, Z_n) through Array_n(X_n, Z_n). As described below, users can navigate the stored images in much the same manner as the user may navigate through an environment using live camera images.

The general operation of one embodiment of inputting images in storage device 20 for transmission to a user will now be described with reference to FIG. 12 and continuing reference to FIG. 11. As shown in step 1210, a cylindrical array 12-1 is situated around the object(s) located in an environment 1100. The view of each camera 14 is transmitted to server 18 in step 1220. Next, in step 1220, the electronic storage device 20 of the server 18 stores the output of each camera 14 at the storage node address associated with that camera 14. Storage of the images may be effectuated serially, from one camera 14 at a time within the array 12, or by simultaneous transmission of the image data from all of the cameras 14 of each array 12. Once the output for each camera 14 of array 12-1 is stored, cylindrical array 12-1 is removed from the environment (step 1240). In step 1250, a determination is made as to the availability of additional cylindrical arrays 12 of differing diameters to those already situated. If additional cylindrical arrays 12 are desired, the process repeats beginning with step 1210. When no additional arrays 12 are available for situating around the environment, the process of inputting images into storage devices 20 is complete (step 1260). At the end of the process, a matrix of addressable stored images exist.

Upon storing all of the outputs associated with the arrays 12-1 through 12-n, a user may navigate through the environment. Navigation is effectuated by accessing the input of the storage nodes by a user interface device 24. In the

present embodiment, the user inputs generally include moving around the environment or object 1100 by moving to the left or right, moving higher or lower along the z-axis, moving through the environment closer or further from the object 1100, or some combination of moving around and through the environment. For example, a user may access the image stored in the node address Array₃(0, 0) to view an object from the camera previously located at coordinate (0, 0) of Array₃. The user may move directly forward, and therefore closer to the object 1100, by accessing the image stored in Array₂(0, 0) and then Array₁(0, 0). To move further away from the object and to the right and up, the user may move from the image stored in node address Array₁(0, 0) and access the images stored in node address Array₂(1, 1), followed by accessing the image stored in node address Array₃(2, 2), and so on. A user may, of course, move among arrays and/or coordinates by any increments changing the point perspective of the environment with each node. Additionally, a user may jump to a particular camera view of the environment. Thus, a user may move throughout the environment in a manner similar to that described above with respect to accessing output of live cameras. This embodiment, however, allows user to access images that are stored in storage nodes as opposed to accessing live cameras. Moreover, this embodiment provides a convenient system and method to allow a user to move forward and backward in an environment.

It should be noted that although each storage node is associated with a camera view identified by its X, Z coordinate of a particular array, other methods of identifying camera views and storage nodes can be used. For example, other coordinate systems, such as those noting angular displacement from a fixed reference point as well as coordinate systems that indicate relative displacement from the current camera node may be used. It should also be understood that the camera arrays 12 may be other shapes other than cylindrical. Moreover, it is not essential, although often advantageous, that the camera arrays 12 surround the entire environment.

It is to be understood that the foregoing user inputs, namely, move clockwise, move counter-clockwise, up, down, closer to the environment, and further from the environment, are merely general descriptions of movement through the environment. Although the present invention is not so limited, in the present preferred embodiment, movement in each of these general directions is further defined based upon the user input. Moreover the output generated by the server to the user may be mixed when moving among adjacent storage nodes associated with environment views (along the x axis, z axis, or among juxtaposed arrays) to generate seamless movement throughout the environment. Mixing may be accomplished by, but are not limited to, the processes described above.

Embodiments Covered

Although the present invention has been described in terms of certain preferred embodiments, other embodiments that are apparent to those of ordinary skill in the art are also intended to be within the scope of this invention. Accordingly, the scope of the present invention is intended to be limited only by the claims appended hereto.

What is claimed is:

1. A telepresence system for providing a first user with a first display of an environment and a second user with a second display of the environment, the system comprising: an array of cameras, each camera having an associated view of the environment and an associated camera

21

output representing the associated view, the array including at least one camera path wherein each path is defined by a series of cameras having progressively different perspectives of the environment;

a first user interface device associated with the first user having first user inputs associated with movement along a first path in the array;

a second user interface device associated with the second user having second user inputs associated with movement along a second path in the array;

at least one processing element coupled to the user interface devices for receiving user inputs, the processing element configured to:

interpret received first user inputs and select outputs of cameras in the first path, mix the outputs of cameras in the first path in accordance with the received first user inputs by sequentially mosaicing the selected outputs of cameras in the first path and

interpret received second inputs and select outputs of cameras in the second path independently of the first inputs, mix the outputs of cameras in the second path in accordance with the received second user inputs by sequentially mosaicing the selected outputs of cameras in the second path, thereby allowing the first user and second user to navigate simultaneously and independently through the array.

2. The system of claim 1 wherein the system further includes a memory storing additional source output, wherein the additional source output includes other than output of the environment and wherein the user inputs include an indication of viewing the additional source output, and the processing element is further configured to mix camera output and the additional source output upon receiving the indication to view the additional source output by mosaicing the camera output and the additional source output.

3. The system of claim 2 wherein the additional source output includes output from the group of outputs including: computer graphic imagery, virtual world imagery, applets, film clips, and animation.

4. A device for providing a user with a display of an environment in response to user inputs, the system comprising:

an array of cameras, each camera having an associated view of the environment and an associated camera output representing the associated view, the array having one or more paths wherein each path is defined by a series of cameras having progressively different perspectives of the environment;

memory storing an additional source output the additional source output including other than output of the environment; and

at least one processing element coupled to the memory for receiving the additional source output, the processing element configured to interpret user inputs and select, based on the user inputs, a camera output to provide to the user, the additional source output to provide to the user, or both a camera view and the additional source output to provide to the user, and the processing element is configured to mix the additional source output with camera output by mosaicing from the camera output to the additional source output.

5. A system for remote seamless viewing of an environment from an array of cameras, each having an output representing an image of the environment, the device comprising:

an interface device having inputs for selecting a path through at least a portion of the array from which to

22

view the environment, the path including a sequence of cameras, each camera in the sequence having a different point perspective and a field of view that overlaps that of an adjacent cameras; and

a display device for sequentially displaying the image from each camera in the sequence by mosaicing the image of a current camera in the sequence to the image of a next camera in the sequence, thereby providing the user a seamless view of the environment.

6. A method for seamless viewing of an environment, the method comprising:

receiving electronically a first image from an array of cameras having progressively different perspectives of the environment, the first image having a first field of view;

receiving electronically a second image from the array, the second image having a second field of the view that overlaps the first field of view;

receiving electronically a third image from the array, the third image having a third field of view that overlaps the second field of view;

mosaicing the first image with the second image and then mosaicing the second image with the third image; and

displaying the first, second, third and mosaic images in sequence to obtain a seamless view through the environment.

7. The method of claim 6 further including selecting an additional source input to be displayed and mosaicing the additional source output with the third image, wherein the additional source output includes other than output of the environment.

8. The method of claim 6 wherein the first, second and third images correspond to first, second, and third cameras, respectively.

9. The method of claim 8 wherein the second camera is adjacent to the first and third cameras.

10. A telepresence system for providing a first user with a first display of an environment and a second user with a second display of the environment, the system comprising:

an array of cameras, each camera having an associated view of the environment and an associated camera output representing the associated view, the array including at least one camera path, wherein each path is defined by a series of cameras having progressively different perspectives of the environment;

a first interface user device associated with the first user having first user inputs associated with movement along a first path in the array;

a second user interface device associated with the second user having second user inputs associated with movement along a second path in the array;

at least one processing element coupled to the user interface devices for receiving user inputs, the processing element configured to:

interpret received first user inputs and select outputs of cameras in the first path, mix the outputs of cameras in the first path in accordance with the received first user inputs by sequentially tweening the selected outputs of cameras in the first path, and

interpret received second inputs and select outputs of cameras in the second path independently of the first inputs, mix the outputs of cameras in the second path in accordance with the received second user inputs by sequentially tweening the selected outputs of cameras in the second path, thereby allowing the first

23

user and second user to navigate simultaneously and independently through the array.

11. The system of claim 10 wherein the system further includes a memory storing additional source output, wherein the additional source output includes other than output of the environment and wherein the user inputs include an indication of viewing the additional source output, and the processing element is further configured to mix camera output and the additional source output upon receiving the indication to view the additional source output by tweening the camera output and the additional source output.

12. The system of claim 11 wherein the additional source output includes output from the group of outputs including: computer graphic imagery, virtual world imagery, applets, film clips, and animation.

13. A device for providing a user with a display of an environment in response to user inputs, the system comprising:

an array of cameras, each camera having an associated view of the environment and an associated camera output representing the associated view, the array having one or more paths wherein each path is defined by a series of cameras having progressively different perspectives of the environment;

memory storing an additional source output, the additional source output including output other than of the environment; and

at least one processing element coupled to the memory for receiving the additional source output, the processing element configured to interpret user inputs and select, based on the user inputs, a camera output to provide to the user, the additional source output to provide to the user, or both a camera view and the additional source output to provide to the user, and the processing element is configured to mix the additional source output with camera output by tweening from the camera output to the additional source output.

14. A system for remote seamless viewing of an environment from an array of cameras, each having an output representing an image of the environment, the device comprising:

an interface device having inputs for selecting a path through at least a portion of the array from which to view the environment, the path including a sequence of cameras, each camera in the sequence having a different point perspective and a field of view that overlaps that of an adjacent cameras; and

a display device for sequentially displaying the image from each camera in the sequence by tweening the image of a current camera in the sequence to the image of a next camera in the sequence, thereby providing the user a seamless view of the environment.

15. A method for seamless viewing of an environment, the method comprising:

receiving electronically a first image from an array of cameras having progressively different perspectives of the environment, the first image having a first field of view;

receiving electronically a second image from the array, the second image having a second field of view that overlaps the first field of view;

receiving electronically a third image from the array, the third image having a third field of view that overlaps the second field of view;

tweening the first image with the second image to obtain a first tweened image and then tweening the second image with the third image to obtain a second tweened image; and

24

displaying the first image, first tweened image, second image, second tweened image, and third image in sequence to obtain a seamless view through the environment.

16. The method of claim 15 further including selecting an additional source output to be displayed and tweening the additional source output with the third image, wherein the additional source output includes other than output of the environment and the displaying includes displaying the additional source output after the third image.

17. The method of claim 15 wherein the first, second and third images correspond to first, second, and third cameras, respectively.

18. The method of claim 17 wherein the second camera is adjacent to the first and third cameras.

19. A method of providing users with real time views of a remote environment, the method comprising:

receiving electronic images of the environment from an array of cameras, the array including at least one camera path through the environment, each path including a series of cameras having progressively different perspectives of the environment;

receiving a first input from a first user interface device associated with a first user, the first input indicating movement along a first path;

receiving a second input from a second user interface device associated with a second user, the second input indicating movement along a second path;

obtaining a first mixed image by mosaicing or tweening, with a first processing element, a first image with a second image in accordance with the first input;

obtaining a second mixed image by mosaicing or tweening, with a second processing element, a third image with a fourth image in accordance with the second input;

providing the first user with the first mixed image in substantially real time, thereby simulating movement along the first path; and

providing the second user with the second mixed image substantially in real time and simultaneous to providing the first user with the first mixed image, thereby independently simulating movement along the second path.

20. The method of claim 19 wherein receiving the first input includes receiving an indication of mosaicing or tweening an additional source output, the method further including obtaining a third mixed output by mosaicing or tweening the second image with the additional source output, wherein the additional source output includes other than output of the environment.

21. The method of claim 19 further including obtaining a third mixed image by mosaicing or tweening the second image with a fifth image in accordance with the first input and providing the first user the third mixed image.

22. A telepresence system for providing a first user with a first display of an environment and a second user with a second display of the environment, the system comprising:

a plurality of removable arrays of cameras, each camera having an associated view of the environment and an associated camera output representing the associated view;

at least one storage device including a plurality of storage nodes wherein the output of each camera is stored in an associated storage node, the storage nodes are accessible to permit at least one path for viewing the environment;

a first user interface device associated with the first user having first user inputs associated with movement

25

along a first path in the environment, the first path including views from cameras in multiple arrays;

a second user interface device associated with the second user having second user inputs associated with movement along a second path in the environment;

at least one processing element coupled to the user interface devices for receiving user inputs including moving up down, clockwise around an environment, counter-clockwise around an environment, forward and backward indicative of movement through the environment, the processing element configured to interpret received first inputs and select outputs of the storage nodes forming the first path, and interpret received second inputs and select outputs of storage nodes forming the second path independently of the first inputs, thereby allowing the first user and second user to navigate simultaneously and independently through the environment, wherein each removable array is situated at different lengths from the environment and the first path includes a first view from a first array and a second view from a second array, thereby allowing the first user to simulate movement forwards and backwards in the environment.

23. The telepresence system of claim 22 wherein the outputs of the cameras are accessible by the processing element.

24. The telepresence system of claim 23 wherein each array is removed after the cameras in the array have transmitted the output to the associated storage node.

25. The telepresence system of claim 24 wherein each array is of cylindrical shape and of a varying diameter.

26. A telepresence system for providing a first user with a first display of an environment and a second user with a second display of the environment, the system comprising:

a plurality of removable arrays of cameras, each camera having an associated view of the environment and an associated camera output representing the associated view, the arrays situated at varying lengths from the environment and including at least one path for viewing the environment, each array capable of being removed after the cameras in the array have transmitted the output to an associated storage node;

at least one storage device including a plurality of storage nodes wherein the output of each camera is stored in an associated storage node, the storage nodes are accessible to permit at least one path for viewing the environment;

26

a first user interface device associated with the first user having first user inputs associated with movement along a first path in the environment;

a second user interface device associated with the second user having second user inputs associated with movement along a second path in the environment;

at least one processing element coupled to the user interface devices for receiving user inputs indicative of movement through the environment, the processing element configured to interpret received first inputs and select outputs of the storage node forming the first path, and interpret received second inputs and select outputs of the storage node forming the second path independently of the first inputs, thereby allowing the first user and second user to navigate simultaneously and independently through the environment, including movement forwards and backwards in the environment.

27. A method of providing a user with views of a remote environment, the method comprising:

receiving electronic images of the environment from a first array of cameras;

removing the first array from the environment;

receiving electronic images of the environment from a second array of cameras, wherein the first and second arrays capture images from different places in the environment;

storing images of the environment from at least the first array for retrieval;

receiving a first input from a first user interface device associated with a first user, the first input indicating movement along a first path, the first path including a first image from the first array and a second image from the second array;

sequentially providing the first user with the first image, and the second image, thereby simulating movement along the first path forward or backward in the environment.

28. The method of claim 27 wherein the arrays are situated concentric in the environment.

29. The method of claim 27 further comprising mixing the first image with the second image to produce a mixed image and sequentially providing the user with the first image, the mixed image and the second image.

30. The method of claim 28 wherein the path is defined by cameras each having a progressively different perspective of the environment.

* * * * *